CHAPTER 3

Systems of Linear Equations

3.1 Introduction

Systems of linear equations play an important and motivating role in the subject of linear algebra. In fact, many problems in linear algebra reduce to finding the solution of a system of linear equations. Thus, the techniques introduced in this chapter will be applicable to abstract ideas introduced later. On the other hand, some of the abstract results will give us new insights into the structure and properties of systems of linear equations.

All our systems of linear equations involve scalars as both coefficients and constants, and such scalars may come from any number field K. There is almost no loss in generality if the reader assumes that all our scalars are real numbers—that is, that they come from the real field $\mathbf{R}$.

3.2 Basic Definitions, Solutions

This section gives basic definitions connected with the solutions of systems of linear equations. The actual algorithms for finding such solutions will be treated later.

Linear Equation and Solutions

A *linear equation* in unknowns $x_1, x_2, \ldots, x_n$ is an equation that can be put in the *standard form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{3.1}$$

where $a_1, a_2, \ldots, a_n$, and b are constants. The constant a_k is called the *coefficient* of x_k, and b is called the *constant term* of the equation.

A solution of the linear equation (3.1) is a list of values for the unknowns or, equivalently, a vector u in K^n, say

$$x_1 = k_1, \quad x_2 = k_2, \quad \ldots, \quad x_n = k_n \qquad \text{or} \qquad u = (k_1, k_2, \ldots, k_n)$$

such that the following statement (obtained by substituting k_i for x_i in the equation) is true:

$$a_1k_1 + a_2k_2 + \cdots + a_nk_n = b$$

In such a case we say that u *satisfies* the equation.

Remark: Equation (3.1) implicitly assumes there is an ordering of the unknowns. In order to avoid subscripts, we will usually use x, y for two unknowns; x, y, z for three unknowns; and x, y, z, t for four unknowns; they will be ordered as shown.

EXAMPLE 3.1 Consider the following linear equation in three unknowns x, y, z:

$$x + 2y - 3z = 6$$

We note that $x = 5, y = 2, z = 1$, or, equivalently, the vector $u = (5, 2, 1)$ is a solution of the equation. That is,

$$5 + 2(2) - 3(1) = 6 \quad \text{or} \quad 5 + 4 - 3 = 6 \quad \text{or} \quad 6 = 6$$

On the other hand, $w = (1, 2, 3)$ is not a solution, because on substitution, we do not get a true statement:

$$1 + 2(2) - 3(3) = 6 \quad \text{or} \quad 1 + 4 - 9 = 6 \quad \text{or} \quad -4 = 6$$

System of Linear Equations

A system of linear equations is a list of linear equations with the same unknowns. In particular, a system of m linear equations $L_1, L_2, \ldots, L_m$ in n unknowns $x_1, x_2, \ldots, x_n$ can be put in the *standard form*

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \tag{3.2}$$

where the a_{ij} and b_i are constants. The number a_{ij} is the *coefficient* of the unknown x_j in the equation L_i, and the number b_i is the *constant* of the equation L_i.

The system (3.2) is called an $m \times n$ (read: m by n) system. It is called a *square system* if $m = n$—that is, if the number m of equations is equal to the number n of unknowns.

The system (3.2) is said to be *homogeneous* if all the constant terms are zero—that is, if $b_1 = 0$, $b_2 = 0, \ldots, b_m = 0$. Otherwise the system is said to be *nonhomogeneous*.

A *solution* (or a *particular solution*) of the system (3.2) is a list of values for the unknowns or, equivalently, a vector u in K^n, which is a solution of each of the equations in the system. The set of all solutions of the system is called the *solution set* or the *general solution* of the system.

EXAMPLE 3.2 Consider the following system of linear equations:

$$\begin{aligned} x_1 + x_2 + 4x_3 + 3x_4 &= 5 \\ 2x_1 + 3x_2 + x_3 - 2x_4 &= 1 \\ x_1 + 2x_2 - 5x_3 + 4x_4 &= 3 \end{aligned}$$

It is a 3×4 system because it has three equations in four unknowns. Determine whether (a) $u = (-8, 6, 1, 1)$ and (b) $v = (-10, 5, 1, 2)$ are solutions of the system.

(a) Substitute the values of u in each equation, obtaining

$$\begin{array}{rcrcr} -8 + 6 + 4(1) + 3(1) = 5 & \text{or} & -8 + 6 + 4 + 3 = 5 & \text{or} & 5 = 5 \\ 2(-8) + 3(6) + 1 - 2(1) = 1 & \text{or} & -16 + 18 + 1 - 2 = 1 & \text{or} & 1 = 1 \\ -8 + 2(6) - 5(1) + 4(1) = 3 & \text{or} & -8 + 12 - 5 + 4 = 3 & \text{or} & 3 = 3 \end{array}$$

Yes, u is a solution of the system because it is a solution of each equation.

(b) Substitute the values of v into each successive equation, obtaining

$$\begin{array}{rcrcr} -10 + 5 + 4(1) + 3(2) = 5 & \text{or} & -10 + 5 + 4 + 6 = 5 & \text{or} & 5 = 5 \\ 2(-10) + 3(5) + 1 - 2(2) = 1 & \text{or} & -20 + 15 + 1 - 4 = 1 & \text{or} & -8 = 1 \end{array}$$

No, v is not a solution of the system, because it is not a solution of the second equation. (We do not need to substitute v into the third equation.)

The system (3.2) of linear equations is said to be *consistent* if it has one or more solutions, and it is said to be *inconsistent* if it has no solution. If the field K of scalars is infinite, such as when K is the real field $\mathbf{R}$ or the complex field $\mathbf{C}$, then we have the following important result.

THEOREM 3.1: Suppose the field K is infinite. Then any system $\mathscr{L}$ of linear equations has (i) a unique solution, (ii) no solution, or (iii) an infinite number of solutions.

This situation is pictured in Fig. 3-1. The three cases have a geometrical description when the system $\mathscr{L}$ consists of two equations in two unknowns (Section 3.4).

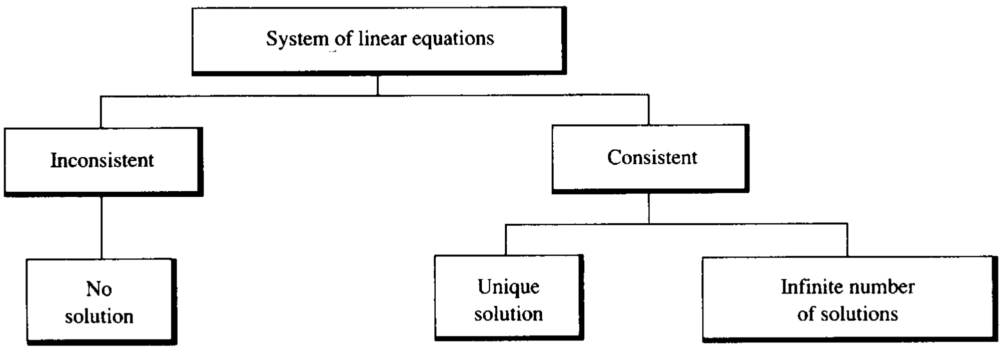


Figure 3-1

Augmented and Coefficient Matrices of a System

Consider again the general system (3.2) of m equations in n unknowns. Such a system has associated with it the following two matrices:

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

The first matrix M is called the *augmented matrix* of the system, and the second matrix A is called the *coefficient matrix*.

The coefficient matrix A is simply the matrix of coefficients, which is the augmented matrix M without the last column of constants. Some texts write $M = [A, B]$ to emphasize the two parts of M, where B denotes the column vector of constants. The augmented matrix M and the coefficient matrix A of the system in Example 3.2 are as follows:

$$M = \begin{bmatrix} 1 & 1 & 4 & 3 & 5 \\ 2 & 3 & 1 & -2 & 1 \\ 1 & 2 & -5 & 4 & 3 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 1 & 1 & 4 & 3 \\ 2 & 3 & 1 & -2 \\ 1 & 2 & -5 & 4 \end{bmatrix}$$

As expected, A consists of all the columns of M except the last, which is the column of constants.

Clearly, a system of linear equations is completely determined by its augmented matrix M, and vice versa. Specifically, each row of M corresponds to an equation of the system, and each column of M corresponds to the coefficients of an unknown, except for the last column, which corresponds to the constants of the system.

Degenerate Linear Equations

A linear equation is said to be *degenerate* if all the coefficients are zero—that is, if it has the form

$$0x_1 + 0x_2 + \cdots + 0x_n = b \tag{3.3}$$

The solution of such an equation depends only on the value of the constant b. Specifically,

(i) If $b \neq 0$, then the equation has no solution.

(ii) If $b = 0$, then every vector $u = (k_1, k_2, \ldots, k_n)$ in K^n is a solution.

The following theorem applies.

THEOREM 3.2: Let $\mathscr{L}$ be a system of linear equations that contains a degenerate equation L, say with constant b.

(i) If $b \neq 0$, then the system $\mathscr{L}$ has no solution.

(ii) If $b = 0$, then L may be deleted from the system without changing the solution set of the system.

Part (i) comes from the fact that the degenerate equation has no solution, so the system has no solution. Part (ii) comes from the fact that every element in K^n is a solution of the degenerate equation.

Leading Unknown in a Nondegenerate Linear Equation

Now let L be a nondegenerate linear equation. This means one or more of the coefficients of L are not zero. By the *leading unknown* of L, we mean the first unknown in L with a nonzero coefficient. For example, x_3 and y are the leading unknowns, respectively, in the equations

$$0x_1 + 0x_2 + 5x_3 + 6x_4 + 0x_5 + 8x_6 = 7 \qquad \text{and} \qquad 0x + 2y - 4z = 5$$

We frequently omit terms with zero coefficients, so the above equations would be written as

$$5x_3 + 6x_4 + 8x_6 = 7 \qquad \text{and} \qquad 2y - 4z = 5$$

In such a case, the leading unknown appears first.

3.3 Equivalent Systems, Elementary Operations

Consider the system (3.2) of m linear equations in n unknowns. Let L be the linear equation obtained by multiplying the m equations by constants $c_1, c_2, \ldots, c_m$, respectively, and then adding the resulting equations. Specifically, let L be the following linear equation:

$$(c_1a_{11} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})x_n = c_1b_1 + \cdots + c_mb_m$$

Then L is called a *linear combination* of the equations in the system. One can easily show (Problem 3.43) that any solution of the system (3.2) is also a solution of the linear combination L.

EXAMPLE 3.3 Let L_1, L_2, L_3 denote, respectively, the three equations in Example 3.2. Let L be the equation obtained by multiplying L_1, L_2, L_3 by $3, -2, 4$, respectively, and then adding. Namely,

$$\begin{array}{rrcl}
3L_1: & 3x_1 + 3x_2 + 12x_3 + \ \ 9x_4 & = & 15 \\
-2L_2: & -4x_1 - 6x_2 - \ \ 2x_3 + \ \ 4x_4 & = & -2 \\
4L_1: & 4x_1 + 8x_2 - 20x_3 + 16x_4 & = & 12 \\
\hline
\text{(Sum) } L: & 3x_1 + 5x_2 - 10x_3 + 29x_4 & = & 25
\end{array}$$

Then L is a linear combination of L_1, L_2, L_3. As expected, the solution $u = (-8, 6, 1, 1)$ of the system is also a solution of L. That is, substituting u in L, we obtain a true statement:

$$3(-8) + 5(6) - 10(1) + 29(1) = 25 \qquad \text{or} \qquad -24 + 30 - 10 + 29 = 25 \qquad \text{or} \qquad 9 = 9$$

The following theorem holds.

THEOREM 3.3: Two systems of linear equations have the same solutions if and only if each equation in each system is a linear combination of the equations in the other system.

Two systems of linear equations are said to be *equivalent* if they have the same solutions. The next subsection shows one way to obtain equivalent systems of linear equations.

Elementary Operations

The following operations on a system of linear equations $L_1, L_2, \ldots, L_m$ are called *elementary operations*.

[E_1] Interchange two of the equations. We indicate that the equations L_i and L_j are interchanged by writing:

$$\text{“Interchange } L_i \text{ and } L_j\text{”} \qquad \text{or} \qquad \text{“}L_i \longleftrightarrow L_j\text{”}$$

[E_2] Replace an equation by a nonzero multiple of itself. We indicate that equation L_i is replaced by kL_i (where $k \neq 0$) by writing

$$\text{“Replace } L_i \text{ by } kL_i\text{”} \qquad \text{or} \qquad \text{“}kL_i \to L_i\text{”}$$

[E_3] Replace an equation by the sum of a multiple of another equation and itself. We indicate that equation L_j is replaced by the sum of kL_i and L_j by writing

$$\text{“Replace } L_j \text{ by } kL_i + L_j\text{”} \qquad \text{or} \qquad \text{“}kL_i + L_j \to L_j\text{”}$$

The arrow $\to$ in [E_2] and [E_3] may be read as "replaces."

The main property of the above elementary operations is contained in the following theorem (proved in Problem 3.45).

THEOREM 3.4: Suppose a system of $\mathscr{M}$ of linear equations is obtained from a system $\mathscr{L}$ of linear equations by a finite sequence of elementary operations. Then $\mathscr{M}$ and $\mathscr{L}$ have the same solutions.

Remark: Sometimes (say to avoid fractions when all the given scalars are integers) we may apply [E_2] and [E_3] in one step; that is, we may apply the following operation:

[E] Replace equation L_j by the sum of kL_i and $k'L_j$ (where $k' \neq 0$), written

$$\text{“Replace } L_j \text{ by } kL_i + k'L_j\text{”} \qquad \text{or} \qquad \text{“}kL_i + k'L_j \to L_j\text{”}$$

We emphasize that in operations [E_3] and [E], only equation L_j is changed.

> Gaussian elimination, our main method for finding the solution of a given system of linear equations, consists of using the above operations to transform a given system into an equivalent system whose solution can be easily obtained.

The details of Gaussian elimination are discussed in subsequent sections.

3.4 Small Square Systems of Linear Equations

This section considers the special case of one equation in one unknown, and two equations in two unknowns. These simple systems are treated separately because their solution sets can be described geometrically, and their properties motivate the general case.

Linear Equation in One Unknown

The following simple basic result is proved in Problem 3.5.

THEOREM 3.5: Consider the linear equation $ax = b$.

(i) If $a \neq 0$, then $x = b/a$ is a unique solution of $ax = b$.

(ii) If $a = 0$, but $b \neq 0$, then $ax = b$ has no solution.

(iii) If $a = 0$ and $b = 0$, then every scalar k is a solution of $ax = b$.

EXAMPLE 3.4 Solve (a) $4x - 1 = x + 6$, (b) $2x - 5 - x = x + 3$, (c) $4 + x - 3 = 2x + 1 - x$.

(a) Rewrite the equation in standard form obtaining $3x = 7$. Then $x = \frac{7}{3}$ is the unique solution [Theorem 3.5(i)].

(b) Rewrite the equation in standard form, obtaining $0x = 8$. The equation has no solution [Theorem 3.5(ii)].

(c) Rewrite the equation in standard form, obtaining $0x = 0$. Then every scalar k is a solution [Theorem 3.5(iii)].

System of Two Linear Equations in Two Unknowns (2×2 System)

Consider a system of two nondegenerate linear equations in two unknowns x and y, which can be put in the standard form

$$\begin{aligned} A_1x + B_1y &= C_1 \\ A_2x + B_2y &= C_2 \end{aligned} \tag{3.4}$$

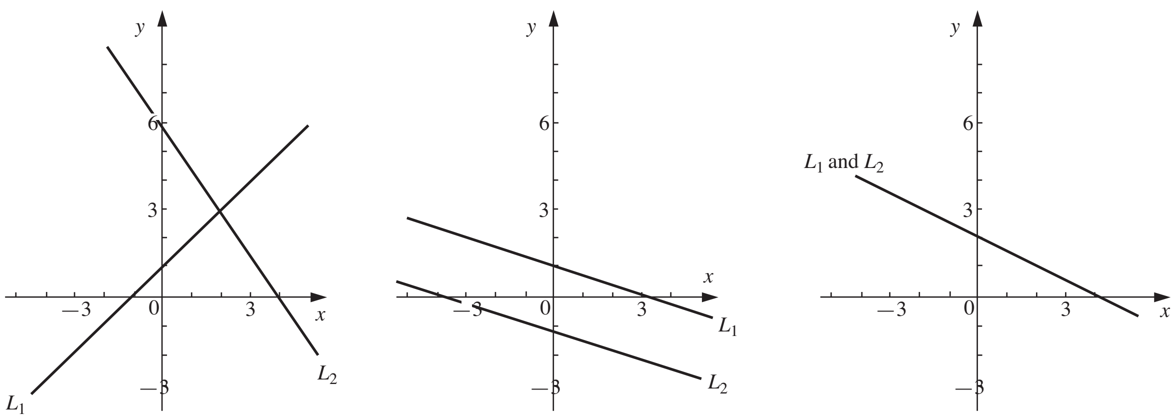
Because the equations are nondegenerate, A_1 and B_1 are not both zero, and A_2 and B_2 are not both zero.

The general solution of the system (3.4) belongs to one of three types as indicated in Fig. 3-1. If **R** is the field of scalars, then the graph of each equation is a line in the plane $\mathbf{R}^2$ and the three types may be described geometrically as pictured in Fig. 3-2. Specifically,

(1) *The system has exactly one solution.*
Here the two lines intersect in one point [Fig. 3-2(a)]. This occurs when the lines have distinct slopes or, equivalently, when the coefficients of x and y are not proportional:

$$\frac{A_1}{A_2} \neq \frac{B_1}{B_2} \qquad \text{or, equivalently,} \qquad A_1B_2 - A_2B_1 \neq 0$$

For example, in Fig. 3-2(a), $1/3 \neq -1/2$.



L_1: $x - y = -1$
L_2: $3x + 2y = 12$
(a)

L_1: $x + 3y = 3$
L_2: $2x + 6y = -8$
(b)

L_1: $x + 2y = 4$
L_2: $2x + 4y = 8$
(c)

Figure 3-2

(2) *The system has no solution.*
Here the two lines are parallel [Fig. 3-2(b)]. This occurs when the lines have the same slopes but different y intercepts, or when

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}$$

For example, in Fig. 3-2(b), $1/2 = 3/6 \neq -3/8$.

(3) *The system has an infinite number of solutions.*
Here the two lines coincide [Fig. 3-2(c)]. This occurs when the lines have the same slopes and same y intercepts, or when the coefficients and constants are proportional,

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$$

For example, in Fig. 3-2(c), $1/2 = 2/4 = 4/8$.

Remark: The following expression and its value is called a *determinant of order two*:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} = A_1B_2 - A_2B_1$$

Determinants will be studied in Chapter 8. Thus, the system (3.4) has a unique solution if and only if the determinant of its coefficients is not zero. (We show later that this statement is true for any square system of linear equations.)

Elimination Algorithm

The solution to system (3.4) can be obtained by the process of elimination, whereby we reduce the system to a single equation in only one unknown. Assuming the system has a unique solution, this elimination algorithm has two parts.

ALGORITHM 3.1: The input consists of two nondegenerate linear equations L_1 and L_2 in two unknowns with a unique solution.

Part A. (Forward Elimination) Multiply each equation by a constant so that the resulting coefficients of one unknown are negatives of each other, and then add the two equations to obtain a new equation L that has only one unknown.

Part B. (Back-Substitution) Solve for the unknown in the new equation L (which contains only one unknown), substitute this value of the unknown into one of the original equations, and then solve to obtain the value of the other unknown.

Part A of Algorithm 3.1 can be applied to any system even if the system does not have a unique solution. In such a case, the new equation L will be degenerate and Part B will not apply.

EXAMPLE 3.5 (Unique Case). Solve the system

$$\begin{aligned} L_1&: \; 2x - 3y = -8 \\ L_2&: \; 3x + 4y = 5 \end{aligned}$$

The unknown x is eliminated from the equations by forming the new equation $L = -3L_1 + 2L_2$. That is, we multiply L_1 by -3 and L_2 by 2 and add the resulting equations as follows:

$$\begin{array}{rr} -3L_1: & -6x + 9y = 24 \\ 2L_2: & 6x + 8y = 10 \\ \hline \text{Addition}: & 17y = 34 \end{array}$$

We now solve the new equation for y, obtaining $y = 2$. We substitute $y = 2$ into one of the original equations, say L_1, and solve for the other unknown x, obtaining

$$2x - 3(2) = -8 \quad \text{or} \quad 2x - 6 = 8 \quad \text{or} \quad 2x = -2 \quad \text{or} \quad x = -1$$

Thus, $x = -1$, $y = 2$, or the pair $u = (-1, 2)$ is the unique solution of the system. The unique solution is expected, because $2/3 \neq -3/4$. [Geometrically, the lines corresponding to the equations intersect at the point $(-1, 2)$.]

EXAMPLE 3.6 (Nonunique Cases)

(a) Solve the system

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = 5 \end{aligned}$$

We eliminated x from the equations by multiplying L_1 by 2 and adding it to L_2—that is, by forming the new equation $L = 2L_1 + L_2$. This yields the degenerate equation

$$0x + 0y = 13$$

which has a nonzero constant $b = 13$. Thus, this equation and the system have no solution. This is expected, because $1/(-2) = -3/6 \neq 4/5$. (Geometrically, the lines corresponding to the equations are parallel.)

(b) Solve the system

$$\begin{aligned} L_1: &\quad x - 3y = 4 \\ L_2: &\quad -2x + 6y = -8 \end{aligned}$$

We eliminated x from the equations by multiplying L_1 by 2 and adding it to L_2—that is, by forming the new equation $L = 2L_1 + L_2$. This yields the degenerate equation

$$0x + 0y = 0$$

where the constant term is also zero. Thus, the system has an infinite number of solutions, which correspond to the solutions of either equation. This is expected, because $1/(-2) = -3/6 = 4/(-8)$. (Geometrically, the lines corresponding to the equations coincide.)

To find the general solution, let $y = a$, and substitute into L_1 to obtain

$$x - 3a = 4 \quad \text{or} \quad x = 3a + 4$$

Thus, the general solution of the system is

$$x = 3a + 4, y = a \quad \text{or} \quad u = (3a + 4,\ a)$$

where a (called a *parameter*) is any scalar.

3.5 Systems in Triangular and Echelon Forms

The main method for solving systems of linear equations, Gaussian elimination, is treated in Section 3.6. Here we consider two simple types of systems of linear equations: systems in triangular form and the more general systems in echelon form.

Triangular Form

Consider the following system of linear equations, which is in *triangular form*:

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 - 2x_4 &= 9 \\ 5x_2 - x_3 + 3x_4 &= 1 \\ 7x_3 - x_4 &= 3 \\ 2x_4 &= 8 \end{aligned}$$

That is, the first unknown x_1 is the leading unknown in the first equation, the second unknown x_2 is the leading unknown in the second equation, and so on. Thus, in particular, the system is square and each leading unknown is *directly* to the right of the leading unknown in the preceding equation.

Such a triangular system always has a unique solution, which may be obtained by *back-substitution*. That is,

(1) First solve the last equation for the last unknown to get $x_4 = 4$.

(2) Then substitute this value $x_4 = 4$ in the next-to-last equation, and solve for the next-to-last unknown x_3 as follows:

$$7x_3 - 4 = 3 \qquad \text{or} \qquad 7x_3 = 7 \qquad \text{or} \qquad x_3 = 1$$

(3) Now substitute $x_3 = 1$ and $x_4 = 4$ in the second equation, and solve for the second unknown x_2 as follows:

$$5x_2 - 1 + 12 = 1 \qquad \text{or} \qquad 5x_2 + 11 = 1 \qquad \text{or} \qquad 5x_2 = -10 \qquad \text{or} \qquad x_2 = -2$$

(4) Finally, substitute $x_2 = -2$, $x_3 = 1$, $x_4 = 4$ in the first equation, and solve for the first unknown x_1 as follows:

$$2x_1 + 6 + 5 - 8 = 9 \qquad \text{or} \qquad 2x_1 + 3 = 9 \qquad \text{or} \qquad 2x_1 = 6 \qquad \text{or} \qquad x_1 = 3$$

Thus, $x_1 = 3$, $x_2 = -2$, $x_3 = 1$, $x_4 = 4$, or, equivalently, the vector $u = (3, -2, 1, 4)$ is the unique solution of the system.

Remark: There is an alternative form for back-substitution (which will be used when solving a system using the matrix format). Namely, after first finding the value of the last unknown, we substitute this value for the last unknown in all the preceding equations before solving for the next-to-last unknown. This yields a triangular system with one less equation and one less unknown. For example, in the above triangular system, we substitute $x_4 = 4$ in all the preceding equations to obtain the triangular system

$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 17 \\ 5x_2 - x_3 &= -1 \\ 7x_3 &= 7 \end{aligned}$$

We then repeat the process using the new last equation. And so on.

Echelon Form, Pivot and Free Variables

The following system of linear equations is said to be in *echelon form*:

$$\begin{aligned} 2x_1 + 6x_2 - x_3 + 4x_4 - 2x_5 &= 15 \\ x_3 + 2x_4 + 2x_5 &= 5 \\ 3x_4 - 9x_5 &= 6 \end{aligned}$$

That is, no equation is degenerate and the leading unknown in each equation other than the first is to the right of the leading unknown in the preceding equation. The leading unknowns in the system, x_1, x_3, x_4, are called *pivot* variables, and the other unknowns, x_2 and x_5, are called *free* variables.

Generally speaking, an *echelon system* or a *system in echelon form* has the following form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots &\\ a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n &= b_r \end{aligned} \tag{3.5}$$

where $1 < j_2 < \cdots < j_r$ and a_{11}, $a_{2j_2}, \ldots, a_{rj_r}$ are not zero. The *pivot* variables are x_1, $x_{j_2}, \ldots, x_{j_r}$. Note that $r \le n$.

The solution set of any echelon system is described in the following theorem (proved in Problem 3.10).

THEOREM 3.6: Consider a system of linear equations in echelon form, say with r equations in n unknowns. There are two cases:

(i) $r = n$. That is, there are as many equations as unknowns (triangular form). Then the system has a unique solution.

(ii) $r < n$. That is, there are more unknowns than equations. Then we can arbitrarily assign values to the $n - r$ free variables and solve uniquely for the r pivot variables, obtaining a solution of the system.

Suppose an echelon system contains more unknowns than equations. Assuming the field K is infinite, the system has an infinite number of solutions, because each of the $n - r$ free variables may be assigned any scalar.

The general solution of a system with free variables may be described in either of two equivalent ways, which we illustrate using the above echelon system where there are $r = 3$ equations and $n = 5$ unknowns. One description is called the ''Parametric Form'' of the solution, and the other description is called the ''Free-Variable Form.''

Parametric Form

Assign arbitrary values, called *parameters*, to the free variables x_2 and x_5, say $x_2 = a$ and $x_5 = b$, and then use back-substitution to obtain values for the pivot variables x_1, x_3, x_5 in terms of the parameters a and b. Specifically,

(1) Substitute $x_5 = b$ in the last equation, and solve for x_4:

$$3x_4 - 9b = 6 \qquad \text{or} \qquad 3x_4 = 6 + 9b \qquad \text{or} \qquad x_4 = 2 + 3b$$

(2) Substitute $x_4 = 2 + 3b$ and $x_5 = b$ into the second equation, and solve for x_3:

$$x_3 + 2(2 + 3b) + 2b = 5 \qquad \text{or} \qquad x_3 + 4 + 8b = 5 \qquad \text{or} \qquad x_3 = 1 - 8b$$

(3) Substitute $x_2 = a$, $x_3 = 1 - 8b$, $x_4 = 2 + 3b$, $x_5 = b$ into the first equation, and solve for x_1:

$$2x_1 + 6a - (1 - 8b) + 4(2 + 3b) - 2b = 15 \qquad \text{or} \qquad x_1 = 4 - 3a - 9b$$

Accordingly, the general solution in *parametric form* is

$$x_1 = 4 - 3a - 9b, \qquad x_2 = a, \qquad x_3 = 1 - 8b, \qquad x_4 = 2 + 3b, \qquad x_5 = b$$

or, equivalently, $v = (4 - 3a - 9b,\ a,\ 1 - 8b,\ 2 + 3b,\ b)$ where a and b are arbitrary numbers.

Free-Variable Form

Use back-substitution to solve for the pivot variables x_1, x_3, x_4 directly in terms of the free variables x_2 and x_5. That is, the last equation gives $x_4 = 2 + 3x_5$. Substitution in the second equation yields $x_3 = 1 - 8x_5$, and then substitution in the first equation yields $x_1 = 4 - 3x_2 - 9x_5$. Accordingly,

$$x_1 = 4 - 3x_2 - 9x_5, \quad x_2 = \text{free variable}, \quad x_3 = 1 - 8x_5, \quad x_4 = 2 + 3x_5, \quad x_5 = \text{free variable}$$

or, equivalently,

$$v = (4 - 3x_2 - 9x_5,\ x_2,\ 1 - 8x_5,\ 2 + 3x_5,\ x_5)$$

is the *free-variable form* for the general solution of the system.

We emphasize that there is no difference between the above two forms of the general solution, and the use of one or the other to represent the general solution is simply a matter of taste.

Remark: A particular solution of the above system can be found by assigning any values to the free variables and then solving for the pivot variables by back-substitution. For example, setting $x_2 = 1$ and $x_5 = 1$, we obtain

$$x_4 = 2 + 3 = 5, \qquad x_3 = 1 - 8 = -7, \qquad x_1 = 4 - 3 - 9 = -8$$

Thus, $u = (-8, 1, 7, 5, 1)$ is the particular solution corresponding to $x_2 = 1$ and $x_5 = 1$.

3.6 Gaussian Elimination

The main method for solving the general system (3.2) of linear equations is called *Gaussian elimination*. It essentially consists of two parts:

Part A. (Forward Elimination) Step-by-step reduction of the system yielding either a degenerate equation with no solution (which indicates the system has no solution) or an equivalent simpler system in triangular or echelon form.

Part B. (Backward Elimination) Step-by-step back-substitution to find the solution of the simpler system.

Part B has already been investigated in Section 3.4. Accordingly, we need only give the algorithm for Part A, which is as follows.

ALGORITHM 3.2 for (Part A): Input: The $m \times n$ system (3.2) of linear equations.

ELIMINATION STEP: Find the first unknown in the system with a nonzero coefficient (which now must be x_1).

(a) Arrange so that $a_{11} \neq 0$. That is, if necessary, interchange equations so that the first unknown x_1 appears with a nonzero coefficient in the first equation.

(b) Use a_{11} *as a pivot* to eliminate x_1 from all equations except the first equation. That is, for $i > 1$:

(1) Set $m = -a_{i1}/a_{11}$; (2) Replace L_i by $mL_1 + L_i$

The system now has the following form:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{2j_2}x_{j_2} + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{mj_2}x_{j_2} + \cdots + a_{mn}x_n &= b_n \end{aligned}$$

where x_1 does not appear in any equation except the first, $a_{11} \neq 0$, and x_{j_2} denotes the first unknown with a nonzero coefficient in any equation other than the first.

(c) Examine each new equation L.

(1) If L has the form $0x_1 + 0x_2 + \cdots + 0x_n = b$ with $b \neq 0$, then

STOP

The system is *inconsistent* and has no solution.

(2) If L has the form $0x_1 + 0x_2 + \cdots + 0x_n = 0$ or if L is a multiple of another equation, then delete L from the system.

RECURSION STEP: Repeat the Elimination Step with each new ‘‘smaller’’ subsystem formed by all the equations excluding the first equation.

OUTPUT: Finally, the system is reduced to triangular or echelon form, or a degenerate equation with no solution is obtained indicating an inconsistent system.

The next remarks refer to the Elimination Step in Algorithm 3.2.

(1) The following number m in (b) is called the *multiplier*:

$$m = -\frac{a_{i1}}{a_{11}} = -\frac{\text{coefficient to be deleted}}{\text{pivot}}$$

(2) One could alternatively apply the following operation in (b):

Replace L_i by $-a_{i1}L_1 + a_{11}L_i$

This would avoid fractions if all the scalars were originally integers.

Gaussian Elimination Example

Here we illustrate in detail Gaussian elimination using the following system of linear equations:

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2x - 4y - 3z &= 8 \\ L_3&: & -3x + 6y + 8z &= -5 \end{aligned}$$

Part A. We use the coefficient 1 of x in the first equation L_1 as the pivot in order to eliminate x from the second equation L_2 and from the third equation L_3. This is accomplished as follows:

(1) Multiply L_1 by the multiplier $m = -2$ and add it to L_2; that is, ‘‘Replace L_2 by $-2L_1 + L_2$.’’

(2) Multiply L_1 by the multiplier $m = 3$ and add it to L_3; that is, ‘‘Replace L_3 by $3L_1 + L_3$.’’

These steps yield

$$\begin{array}{rrcrr} (-2)L_1: & -2x + 6y + 4z = -12 & \qquad & 3L_1: & 3x - 9y - 6z = 18 \\ L_2: & 2x - 4y - 3z = 8 & & L_3: & -3x + 6y + 8z = -5 \\ \hline \text{New } L_2: & 2y + z = -4 & & \text{New } L_3: & -3y + 2z = 13 \end{array}$$

Thus, the original system is replaced by the following system:

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2y + z &= -4 \\ L_3&: & -3y + 2z &= 13 \end{aligned}$$

(Note that the equations L_2 and L_3 form a subsystem with one less equation and one less unknown than the original system.)

Next we use the coefficient 2 of y in the (new) second equation L_2 as the pivot in order to eliminate y from the (new) third equation L_3. This is accomplished as follows:

(3) Multiply L_2 by the multiplier $m = \frac{3}{2}$ and add it to L_3; that is, ‘‘Replace L_3 by $\frac{3}{2}L_2 + L_3$.’’ (Alternately, ‘‘Replace L_3 by $3L_2 + 2L_3$,’’ which will avoid fractions.)

This step yields

$$\begin{array}{rrcrr} \frac{3}{2}L_2: & 3y + \frac{3}{2}z = -6 & & 3L_2: & 6y + 3z = -12 \\ L_3: & -3y + 2z = 13 & \text{or} & 2L_3: & -6y + 4z = 26 \\ \hline \text{New } L_3: & \frac{7}{2}z = 7 & & \text{New } L_3: & 7z = 14 \end{array}$$

Thus, our system is replaced by the following system:

$$\begin{aligned} L_1&: & x - 3y - 2z &= 6 \\ L_2&: & 2y + z &= -4 \\ L_3&: & 7z &= 14 \quad (\text{or } \tfrac{7}{2}z = 7) \end{aligned}$$

The system is now in triangular form, so Part A is completed.

Part B. The values for the unknowns are obtained in reverse order, z, y, x, by back-substitution. Specifically,

(1) Solve for z in L_3 to get $z = 2$.

(2) Substitute $z = 2$ in L_2, and solve for y to get $y = -3$.

(3) Substitute $y = -3$ and $z = 2$ in L_1, and solve for x to get $x = 1$.

Thus, the solution of the triangular system and hence the original system is as follows:

$$x = 1, \quad y = -3, \quad z = 2 \qquad \text{or, equivalently,} \qquad u = (1, -3, 2).$$

Condensed Format

The Gaussian elimination algorithm involves rewriting systems of linear equations. Sometimes we can avoid excessive recopying of some of the equations by adopting a ‘‘condensed format.’’ This format for the solution of the above system follows:

Number	Equation	Operation
(1)	$x - 3y - 2z = 6$	
(2)	$2x - 4y - 3z = 8$	
(3)	$-3x + 6y + 8z = -5$	
$(2')$	$2y + z = -4$	Replace L_2 by $-2L_1 + L_2$
$(3')$	$-3y + 2z = 13$	Replace L_3 by $3L_1 + L_3$
$(3'')$	$7z = 14$	Replace L_3 by $3L_2 + 2L_3$

That is, first we write down the number of each of the original equations. As we apply the Gaussian elimination algorithm to the system, we only write down the new equations, and we label each new equation using the same number as the original corresponding equation, but with an added prime. (After each new equation, we will indicate, for instructional purposes, the elementary operation that yielded the new equation.)

The system in triangular form consists of equations (1), $(2')$, and $(3'')$, the numbers with the largest number of primes. Applying back-substitution to these equations again yields $x = 1$, $y = -3$, $z = 2$.

Remark: If two equations need to be interchanged, say to obtain a nonzero coefficient as a pivot, then this is easily accomplished in the format by simply renumbering the two equations rather than changing their positions.

EXAMPLE 3.7 Solve the following system:

$$\begin{aligned} x + 2y - 3z &= 1 \\ 2x + 5y - 8z &= 4 \\ 3x + 8y - 13z &= 7 \end{aligned}$$

We solve the system by Gaussian elimination.

Part A. (Forward Elimination) We use the coefficient 1 of x in the first equation L_1 as the pivot in order to eliminate x from the second equation L_2 and from the third equation L_3. This is accomplished as follows:

(1) Multiply L_1 by the multiplier $m = -2$ and add it to L_2; that is, ‘‘Replace L_2 by $-2L_1 + L_2$.’’

(2) Multiply L_1 by the multiplier $m = -3$ and add it to L_3; that is, ‘‘Replace L_3 by $-3L_1 + L_3$.’’

The two steps yield

$$\begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \\ 2y - 4z &= 4 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + 2y - 3z &= 1 \\ y - 2z &= 2 \end{aligned}$$

(The third equation is deleted, because it is a multiple of the second equation.) The system is now in echelon form with free variable z.

Part B. (Backward Elimination) To obtain the general solution, let the free variable $z = a$, and solve for x and y by back-substitution. Substitute $z = a$ in the second equation to obtain $y = 2 + 2a$. Then substitute $z = a$ and $y = 2 + 2a$ into the first equation to obtain

$$x + 2(2 + 2a) - 3a = 1 \qquad \text{or} \qquad x + 4 + 4a - 3a = 1 \qquad \text{or} \qquad x = -3 - a$$

Thus, the following is the general solution where a is a parameter:

$$x = -3 - a, \quad y = 2 + 2a, \quad z = a \qquad \text{or} \qquad u = (-3 - a,\ 2 + 2a,\ a)$$

EXAMPLE 3.8 Solve the following system:

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\ 2x_1 + 8x_2 - x_3 + 9x_4 &= 9 \\ 3x_1 + 5x_2 - 12x_3 + 17x_4 &= 7 \end{aligned}$$

We use Gaussian elimination.

Part A. (Forward Elimination) We use the coefficient 1 of x_1 in the first equation L_1 as the pivot in order to eliminate x_1 from the second equation L_2 and from the third equation L_3. This is accomplished by the following operations:

(1) "Replace L_2 by $-2L_1 + L_2$" and (2) "Replace L_3 by $-3L_1 + L_3$"

These yield:

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\ 2x_2 + 3x_3 - x_4 &= 1 \\ -4x_2 - 6x_3 + 2x_4 &= -5 \end{aligned}$$

We now use the coefficient 2 of x_2 in the second equation L_2 as the pivot and the multiplier $m = 2$ in order to eliminate x_2 from the third equation L_3. This is accomplished by the operation "Replace L_3 by $2L_2 + L_3$," which then yields the degenerate equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -3$$

This equation and, hence, the original system have no solution:

DO NOT CONTINUE

Remark 1: As in the above examples, Part A of Gaussian elimination tells us whether or not the system has a solution—that is, whether or not the system is consistent. Accordingly, Part B need never be applied when a system has no solution.

Remark 2: If a system of linear equations has more than four unknowns and four equations, then it may be more convenient to use the matrix format for solving the system. This matrix format is discussed later.

3.7 Echelon Matrices, Row Canonical Form, Row Equivalence

One way to solve a system of linear equations is by working with its augmented matrix M rather than the system itself. This section introduces the necessary matrix concepts for such a discussion. These concepts, such as echelon matrices and elementary row operations, are also of independent interest.

Echelon Matrices

A matrix A is called an *echelon matrix*, or is said to be in *echelon form*, if the following two conditions hold (where a *leading nonzero element* of a row of A is the first nonzero element in the row):

(1) All zero rows, if any, are at the bottom of the matrix.

(2) Each leading nonzero entry in a row is to the right of the leading nonzero entry in the preceding row.

That is, $A = [a_{ij}]$ is an echelon matrix if there exist nonzero entries

$$a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}, \qquad \text{where} \quad j_1 < j_2 < \cdots < j_r$$

with the property that

$$a_{ij} = 0 \qquad \text{for} \quad \begin{cases} \text{(i) } i \le r, \quad j < j_i \\ \text{(ii) } i > r \end{cases}$$

The entries $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$, which are the leading nonzero elements in their respective rows, are called the *pivots* of the echelon matrix.

EXAMPLE 3.9 The following is an echelon matrix whose pivots have been circled:

$$A = \begin{bmatrix} 0 & 2 & 3 & 4 & 5 & 9 & 0 & 7 \\ 0 & 0 & 0 & 3 & 4 & 1 & 2 & 5 \\ 0 & 0 & 0 & 0 & 0 & 5 & 7 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 8 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Observe that the pivots are in columns C_2, C_4, C_6, C_7, and each is to the right of the one above. Using the above notation, the pivots are

$$a_{1j_1} = 2, \qquad a_{2j_2} = 3, \qquad a_{3j_3} = 5, \qquad a_{4j_4} = 8$$

where $j_1 = 2$, $j_2 = 4$, $j_3 = 6$, $j_4 = 7$. Here $r = 4$.

Row Canonical Form

A matrix A is said to be in *row canonical form* (or *row-reduced echelon form*) if it is an echelon matrix—that is, if it satisfies the above properties (1) and (2), and if it satisfies the following additional two properties:

(3) Each pivot (leading nonzero entry) is equal to 1.

(4) Each pivot is the only nonzero entry in its column.

The major difference between an echelon matrix and a matrix in row canonical form is that in an echelon matrix there must be zeros below the pivots [Properties (1) and (2)], but in a matrix in row canonical form, each pivot must also equal 1 [Property (3)] and there must also be zeros above the pivots [Property (4)].

The zero matrix 0 of any size and the identity matrix I of any size are important special examples of matrices in row canonical form.

EXAMPLE 3.10

The following are echelon matrices whose pivots have been circled:

$$\begin{bmatrix} 2 & 3 & 2 & 0 & 4 & 5 & -6 \\ 0 & 0 & 0 & 1 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 0 & 1 & 3 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

The third matrix is also an example of a matrix in row canonical form. The second matrix is not in row canonical form, because it does not satisfy property (4); that is, there is a nonzero entry above the second pivot in the third column. The first matrix is not in row canonical form, because it satisfies neither property (3) nor property (4); that is, some pivots are not equal to 1 and there are nonzero entries above the pivots.

Elementary Row Operations

Suppose A is a matrix with rows $R_1, R_2, \ldots, R_m$. The following operations on A are called *elementary row operations*.

[E_1] (Row Interchange): Interchange rows R_i and R_j. This may be written as

$$\text{"Interchange } R_i \text{ and } R_j\text{"} \qquad \text{or} \qquad \text{"}R_i \longleftrightarrow R_j\text{"}$$

[E_2] (Row Scaling): Replace row R_i by a nonzero multiple kR_i of itself. This may be written as

$$\text{"Replace } R_i \text{ by } kR_i \ (k \neq 0)\text{"} \qquad \text{or} \qquad \text{"}kR_i \to R_i\text{"}$$

[E_3] (Row Addition): Replace row R_j by the sum of a multiple kR_i of a row R_i and itself. This may be written as

$$\text{"Replace } R_j \text{ by } kR_i + R_j\text{"} \qquad \text{or} \qquad \text{"}kR_i + R_j \to R_j\text{"}$$

The arrow $\to$ in E_2 and E_3 may be read as "replaces."

Sometimes (say to avoid fractions when all the given scalars are integers) we may apply [E_2] and [E_3] in one step; that is, we may apply the following operation:

[E] Replace R_j by the sum of a multiple kR_i of a row R_i and a nonzero multiple $k'R_j$ of itself. This may be written as

$$\text{"Replace } R_j \text{ by } kR_i + k'R_j \ (k' \neq 0)\text{"} \qquad \text{or} \qquad \text{"}kR_i + k'R_j \to R_j\text{"}$$

We emphasize that in operations [E_3] and [E] only row R_j is changed.

Row Equivalence, Rank of a Matrix

A matrix A is said to be *row equivalent* to a matrix B, written

$A \sim B$

if B can be obtained from A by a sequence of elementary row operations. In the case that B is also an echelon matrix, B is called an *echelon form* of A.

The following are two basic results on row equivalence.

THEOREM 3.7: Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are row equivalent echelon matrices with respective pivot entries

$$a_{1j_1}, a_{2j_2}, \ldots a_{rj_r} \qquad \text{and} \qquad b_{1k_1}, b_{2k_2}, \ldots b_{sk_s}$$

Then A and B have the same number of nonzero rows—that is, $r = s$—and the pivot entries are in the same positions—that is, $j_1 = k_1, \ j_2 = k_2, \ \ldots, \ j_r = k_r$.

THEOREM 3.8: Every matrix A is row equivalent to a unique matrix in row canonical form.

The proofs of the above theorems will be postponed to Chapter 4. The unique matrix in Theorem 3.8 is called the *row canonical form* of A.

Using the above theorems, we can now give our first definition of the rank of a matrix.

DEFINITION: The *rank* of a matrix A, written $\text{rank}(A)$, is equal to the number of pivots in an echelon form of A.

> The rank is a very important property of a matrix and, depending on the context in which the matrix is used, it will be defined in many different ways. Of course, all the definitions lead to the same number.

The next section gives the matrix format of Gaussian elimination, which finds an echelon form of any matrix A (and hence the rank of A), and also finds the row canonical form of A.

One can show that row equivalence is an *equivalence relation*. That is,

(1) $A \sim A$ for any matrix A.
(2) If $A \sim B$, then $B \sim A$.
(3) If $A \sim B$ and $B \sim C$, then $A \sim C$.

Property (2) comes from the fact that each elementary row operation has an inverse operation of the same type. Namely,

(i) ‘‘Interchange R_i and R_j’’ is its own inverse.
(ii) ‘‘Replace R_i by kR_i’’ and ‘‘Replace R_i by $(1/k)R_i$’’ are inverses.
(iii) ‘‘Replace R_j by $kR_i + R_j$’’ and ‘‘Replace R_j by $-kR_i + R_j$’’ are inverses.

There is a similar result for operation [E] (Problem 3.73).

3.8 Gaussian Elimination, Matrix Formulation

This section gives two matrix algorithms that accomplish the following:

(1) Algorithm 3.3 transforms any matrix A into an echelon form.
(2) Algorithm 3.4 transforms the echelon matrix into its row canonical form.

These algorithms, which use the elementary row operations, are simply restatements of *Gaussian elimination* as applied to matrices rather than to linear equations. (The term ‘‘row reduce’’ or simply ‘‘reduce’’ will mean to transform a matrix by the elementary row operations.)

ALGORITHM 3.3 (Forward Elimination): The input is any matrix A. (The algorithm puts 0’s below each pivot, working from the ‘‘top-down.’’) The output is an echelon form of A.

Step 1. Find the first column with a nonzero entry. Let j_1 denote this column.

(a) Arrange so that $a_{1j_1} \neq 0$. That is, if necessary, interchange rows so that a nonzero entry appears in the first row in column j_1.

(b) Use a_{1j_1} as a pivot to obtain 0’s below a_{1j_1}.

Specifically, for $i > 1$:

(1) Set $m = -a_{ij_1}/a_{1j_1}$; (2) Replace R_i by $mR_1 + R_i$

[That is, apply the operation $-(a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$.]

Step 2. Repeat Step 1 with the submatrix formed by all the rows excluding the first row. Here we let j_2 denote the first column in the subsystem with a nonzero entry. Hence, at the end of Step 2, we have $a_{2j_2} \neq 0$.

Steps 3 to *r*. Continue the above process until a submatrix has only zero rows.

We emphasize that at the end of the algorithm, the pivots will be

$$a_{1j_1},\ a_{2j_2},\ \ldots,\ a_{rj_r}$$

where r denotes the number of nonzero rows in the final echelon matrix.

Remark 1: The following number m in Step 1(b) is called the *multiplier*:

$$m = -\frac{a_{ij_1}}{a_{1j_1}} = -\frac{\text{entry to be deleted}}{\text{pivot}}$$

Remark 2: One could replace the operation in Step 1(b) by the following which would avoid fractions if all the scalars were originally integers.

Replace R_i by $-a_{ij_1}R_1 + a_{1j_1}R_i$.

ALGORITHM 3.4 (Backward Elimination): The input is a matrix $A = [a_{ij}]$ in echelon form with pivot entries

$$a_{1j_1}, \quad a_{2j_2}, \quad \ldots, \quad a_{rj_r}$$

The output is the row canonical form of A.

Step 1. (a) (Use row scaling so the last pivot equals 1.) Multiply the last nonzero row R_r by $1/a_{rj_r}$.

(b) (Use $a_{rj_r} = 1$ to obtain 0's above the pivot.) For $i = r-1, \quad r-2, \quad \ldots, \quad 2, \quad 1$:

$$(1)\ \text{Set } m = -a_{ij_r}; \qquad (2)\ \text{Replace } R_i \text{ by } mR_r + R_i$$

(That is, apply the operations $-a_{ij_r}R_r + R_i \to R_i$.)

Steps 2 to *r*−1. Repeat Step 1 for rows $R_{r-1}, R_{r-2}, \ldots, R_2$.

Step *r*. (Use row scaling so the first pivot equals 1.) Multiply R_1 by $1/a_{1j_1}$.

There is an alternative form of Algorithm 3.4, which we describe here in words. The formal description of this algorithm is left to the reader as a supplementary problem.

ALTERNATIVE ALGORITHM 3.4 Puts 0's above the pivots row by row from the bottom up (rather than column by column from right to left).

The alternative algorithm, when applied to an augmented matrix M of a system of linear equations, is essentially the same as solving for the pivot unknowns one after the other from the bottom up.

Remark: We emphasize that Gaussian elimination is a two-stage process. Specifically,

Stage A (Algorithm 3.3). Puts 0's below each pivot, working from the top row R_1 down.

Stage B (Algorithm 3.4). Puts 0's above each pivot, working from the bottom row R_r up.

There is another algorithm, called *Gauss–Jordan*, that also row reduces a matrix to its row canonical form. The difference is that Gauss–Jordan puts 0's both below and above each pivot as it works its way from the top row R_1 down. Although Gauss–Jordan may be easier to state and understand, it is much less efficient than the two-stage Gaussian elimination algorithm.

EXAMPLE 3.11 Consider the matrix $A = \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 2 & 4 & -4 & 6 & 10 \\ 3 & 6 & -6 & 9 & 13 \end{bmatrix}$.

(a) Use Algorithm 3.3 to reduce A to an echelon form.

(b) Use Algorithm 3.4 to further reduce A to its row canonical form.

(a) First use $a_{11} = 1$ as a pivot to obtain 0's below a_{11}; that is, apply the operations ''Replace R_2 by $-2R_1 + R_2$'' and ''Replace R_3 by $-3R_1 + R_3$.'' Then use $a_{23} = 2$ as a pivot to obtain 0 below a_{23}; that is, apply the operation ''Replace R_3 by $-\frac{3}{2}R_2 + R_3$.'' This yields

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

The matrix is now in echelon form.

(b) Multiply R_3 by $-\frac{1}{2}$ so the pivot entry $a_{35} = 1$, and then use $a_{35} = 1$ as a pivot to obtain 0's above it by the operations "Replace R_2 by $-6R_3 + R_2$" and then "Replace R_1 by $-2R_3 + R_1$." This yields

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 2 \\ 0 & 0 & 2 & 4 & 6 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Multiply R_2 by $\frac{1}{2}$ so the pivot entry $a_{23} = 1$, and then use $a_{23} = 1$ as a pivot to obtain 0's above it by the operation "Replace R_1 by $3R_2 + R_1$." This yields

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 7 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The last matrix is the row canonical form of A.

Application to Systems of Linear Equations

One way to solve a system of linear equations is by working with its augmented matrix M rather than the equations themselves. Specifically, we reduce M to echelon form (which tells us whether the system has a solution), and then further reduce M to its row canonical form (which essentially gives the solution of the original system of linear equations). The justification for this process comes from the following facts:

(1) Any elementary row operation on the augmented matrix M of the system is equivalent to applying the corresponding operation on the system itself.

(2) The system has a solution if and only if the echelon form of the augmented matrix M does not have a row of the form $(0, 0, \ldots, 0, b)$ with $b \neq 0$.

(3) In the row canonical form of the augmented matrix M (excluding zero rows), the coefficient of each basic variable is a pivot entry equal to 1, and it is the only nonzero entry in its respective column; hence, the free-variable form of the solution of the system of linear equations is obtained by simply transferring the free variables to the other side.

This process is illustrated below.

EXAMPLE 3.12 Solve each of the following systems:

$$\begin{array}{c} x_1 + x_2 - 2x_3 + 4x_4 = 5 \\ 2x_1 + 2x_2 - 3x_3 + x_4 = 3 \\ 3x_1 + 3x_2 - 4x_3 - 2x_4 = 1 \\ \text{(a)} \end{array} \qquad \begin{array}{c} x_1 + x_2 - 2x_3 + 3x_4 = 4 \\ 2x_1 + 3x_2 + 3x_3 - x_4 = 3 \\ 5x_1 + 7x_2 + 4x_3 + x_4 = 5 \\ \text{(b)} \end{array} \qquad \begin{array}{c} x + 2y + z = 3 \\ 2x + 5y - z = -4 \\ 3x - 2y - z = 5 \\ \text{(c)} \end{array}$$

(a) Reduce its augmented matrix M to echelon form and then to row canonical form as follows:

$$M = \begin{bmatrix} 1 & 1 & -2 & 4 & 5 \\ 2 & 2 & -3 & 1 & 3 \\ 3 & 3 & -4 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 4 & 5 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 2 & -14 & -14 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 & -10 & -9 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Rewrite the row canonical form in terms of a system of linear equations to obtain the free variable form of the solution. That is,

$$\begin{aligned} x_1 + x_2 \quad - 10x_4 &= -9 \\ x_3 - 7x_4 &= -7 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x_1 &= -9 - x_2 + 10x_4 \\ x_3 &= -7 + 7x_4 \end{aligned}$$

(The zero row is omitted in the solution.) Observe that x_1 and x_3 are the pivot variables, and x_2 and x_4 are the free variables.

(b) First reduce its augmented matrix M to echelon form as follows:

$$M = \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 2 & 3 & 3 & -1 & 3 \\ 5 & 7 & 4 & 1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 2 & 14 & -14 & -15 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 & 3 & 4 \\ 0 & 1 & 7 & -7 & -5 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

There is no need to continue to find the row canonical form of M, because the echelon form already tells us that the system has no solution. Specifically, the third row of the echelon matrix corresponds to the degenerate equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = -5$$

which has no solution. Thus, the system has no solution.

(c) Reduce its augmented matrix M to echelon form and then to row canonical form as follows:

$$M = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 5 & -1 & -4 \\ 3 & -2 & -1 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & -8 & -4 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & -28 & -84 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 1 & -3 & -10 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Thus, the system has the unique solution $x = 2$, $y = -1$, $z = 3$, or, equivalently, the vector $u = (2, -1, 3)$. We note that the echelon form of M already indicated that the solution was unique, because it corresponded to a triangular system.

Application to Existence and Uniqueness Theorems

This subsection gives theoretical conditions for the existence and uniqueness of a solution of a system of linear equations using the notion of the rank of a matrix.

THEOREM 3.9: Consider a system of linear equations in n unknowns with augmented matrix $M = [A, B]$. Then,

(a) The system has a solution if and only if $\operatorname{rank}(A) = \operatorname{rank}(M)$.

(b) The solution is unique if and only if $\operatorname{rank}(A) = \operatorname{rank}(M) = n$.

Proof of (a). The system has a solution if and only if an echelon form of $M = [A, B]$ does not have a row of the form

$$(0, 0, \ldots, 0, b), \qquad \text{with} \quad b \neq 0$$

If an echelon form of M does have such a row, then b is a pivot of M but not of A, and hence, $\operatorname{rank}(M) > \operatorname{rank}(A)$. Otherwise, the echelon forms of A and M have the same pivots, and hence, $\operatorname{rank}(A) = \operatorname{rank}(M)$. This proves (a).

Proof of (b). The system has a unique solution if and only if an echelon form has no free variable. This means there is a pivot for each unknown. Accordingly, $n = \operatorname{rank}(A) = \operatorname{rank}(M)$. This proves (b).

The above proof uses the fact (Problem 3.74) that an echelon form of the augmented matrix $M = [A, B]$ also automatically yields an echelon form of A.

3.9 Matrix Equation of a System of Linear Equations

The general system (3.2) of m linear equations in n unknowns is equivalent to the matrix equation

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \qquad \text{or} \qquad AX = B$$

where $A = [a_{ij}]$ is the coefficient matrix, $X = [x_j]$ is the column vector of unknowns, and $B = [b_i]$ is the column vector of constants. (Some texts write $Ax = b$ rather than $AX = B$, in order to emphasize that x and b are simply column vectors.)

The statement that the system of linear equations and the matrix equation are equivalent means that any vector solution of the system is a solution of the matrix equation, and vice versa.

EXAMPLE 3.13 The following system of linear equations and matrix equation are equivalent:

$$\begin{aligned} x_1 + 2x_2 - 4x_3 + 7x_4 &= 4 \\ 3x_1 - 5x_2 + 6x_3 - 8x_4 &= 8 \\ 4x_1 - 3x_2 - 2x_3 + 6x_4 &= 11 \end{aligned} \qquad \text{and} \qquad \begin{bmatrix} 1 & 2 & -4 & 7 \\ 3 & -5 & 6 & -8 \\ 4 & -3 & -2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 11 \end{bmatrix}$$

We note that $x_1 = 3,\ x_2 = 1,\ x_3 = 2,\ x_4 = 1,$ or, in other words, the vector $u = [3, 1, 2, 1]$ is a solution of the system. Thus, the (column) vector u is also a solution of the matrix equation.

The matrix form $AX = B$ of a system of linear equations is notationally very convenient when discussing and proving properties of systems of linear equations. This is illustrated with our first theorem (described in Fig. 3-1), which we restate for easy reference.

THEOREM 3.1: Suppose the field K is infinite. Then the system $AX = B$ has: (a) a unique solution, (b) no solution, or (c) an infinite number of solutions.

Proof. It suffices to show that if $AX = B$ has more than one solution, then it has infinitely many. Suppose u and v are distinct solutions of $AX = B$; that is, $Au = B$ and $Av = B$. Then, for any $k \in K$,

$$A[u + k(u - v)] = Au + k(Au - Av) = B + k(B - B) = B$$

Thus, for each $k \in K$, the vector $u + k(u - v)$ is a solution of $AX = B$. Because all such solutions are distinct (Problem 3.47), $AX = B$ has an infinite number of solutions.

Observe that the above theorem is true when K is the real field $\mathbf{R}$ (or the complex field $\mathbf{C}$). Section 3.3 shows that the theorem has a geometrical description when the system consists of two equations in two unknowns, where each equation represents a line in $\mathbf{R}^2$. The theorem also has a geometrical description when the system consists of three nondegenerate equations in three unknowns, where the three equations correspond to planes H_1, H_2, H_3 in $\mathbf{R}^3$. That is,

(a) *Unique solution:* Here the three planes intersect in exactly one point.

(b) *No solution:* Here the planes may intersect pairwise but with no common point of intersection, or two of the planes may be parallel.

(c) *Infinite number of solutions:* Here the three planes may intersect in a line (one free variable), or they may coincide (two free variables).

These three cases are pictured in Fig. 3-3.

Matrix Equation of a Square System of Linear Equations

A system $AX = B$ of linear equations is square if and only if the matrix A of coefficients is square. In such a case, we have the following important result.

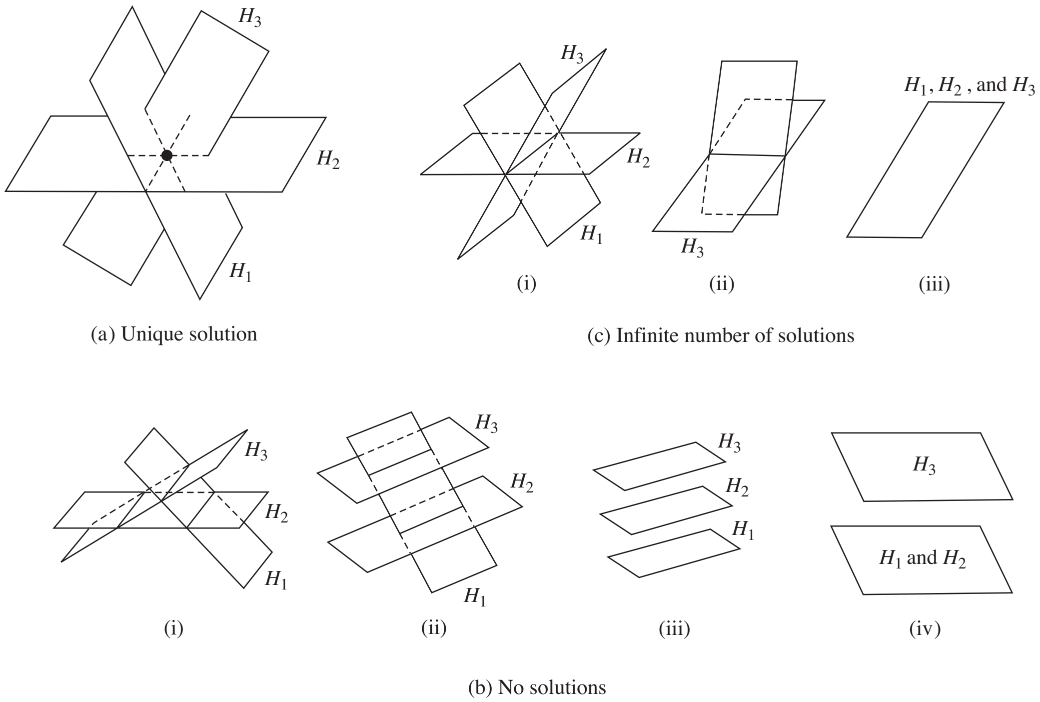


Figure 3-3

THEOREM 3.10: A square system $AX = B$ of linear equations has a unique solution if and only if the matrix A is invertible. In such a case, $A^{-1}B$ is the unique solution of the system.

We only prove here that if A is invertible, then $A^{-1}B$ is a unique solution. If A is invertible, then

$$A(A^{-1}B) = (AA^{-1})B = IB = B$$

and hence, $A^{-1}B$ is a solution. Now suppose v is any solution, so $Av = B$. Then

$$v = Iv = (A^{-1}A)v = A^{-1}(Av) = A^{-1}B$$

Thus, the solution $A^{-1}B$ is unique.

EXAMPLE 3.14 Consider the following system of linear equations, whose coefficient matrix A and inverse A^{-1} are also given:

$$\begin{aligned} x + 2y + 3z &= 1 \\ x + 3y + 6z &= 3, \\ 2x + 6y + 13z &= 5 \end{aligned} \qquad A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 6 \\ 2 & 6 & 13 \end{bmatrix}, \qquad A^{-1} = \begin{bmatrix} 3 & -8 & 3 \\ -1 & 7 & -3 \\ 0 & -2 & 1 \end{bmatrix}$$

By Theorem 3.10, the unique solution of the system is

$$A^{-1}B = \begin{bmatrix} 3 & -8 & 3 \\ -1 & 7 & -3 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -6 \\ 5 \\ -1 \end{bmatrix}$$

That is, $x = -6$, $y = 5$, $z = -1$.

Remark: We emphasize that Theorem 3.10 does not usually help us to find the solution of a square system. That is, finding the inverse of a coefficient matrix A is not usually any easier than solving the system directly. Thus, unless we are given the inverse of a coefficient matrix A, as in Example 3.14, we usually solve a square system by Gaussian elimination (or some iterative method whose discussion lies beyond the scope of this text).

3.10 Systems of Linear Equations and Linear Combinations of Vectors

The general system (3.2) of linear equations may be rewritten as the following vector equation:

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_m \end{bmatrix}$$

Recall that a vector v in K^n is said to be a *linear combination* of vectors $u_1, u_2, \ldots, u_m$ in K^n if there exist scalars $a_1, a_2, \ldots, a_m$ in K such that

$$v = a_1 u_1 + a_2 u_2 + \cdots + a_m u_m$$

Accordingly, the general system (3.2) of linear equations and the above equivalent vector equation have a solution if and only if the column vector of constants is a linear combination of the columns of the coefficient matrix. We state this observation formally.

THEOREM 3.11: A system $AX = B$ of linear equations has a solution if and only if B is a linear combination of the columns of the coefficient matrix A.

Thus, the answer to the problem of expressing a given vector v in K^n as a linear combination of vectors $u_1, u_2, \ldots, u_m$ in K^n reduces to solving a system of linear equations.

Linear Combination Example

Suppose we want to write the vector $v = (1, -2, 5)$ as a linear combination of the vectors

$$u_1 = (1, 1, 1), \qquad u_2 = (1, 2, 3), \qquad u_3 = (2, -1, 1)$$

First we write $v = xu_1 + yu_2 + zu_3$ with unknowns x, y, z, and then we find the equivalent system of linear equations which we solve. Specifically, we first write

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + z \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \tag{*}$$

Then

$$\begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix} = \begin{bmatrix} x \\ x \\ x \end{bmatrix} + \begin{bmatrix} y \\ 2y \\ 3y \end{bmatrix} + \begin{bmatrix} 2z \\ -z \\ z \end{bmatrix} = \begin{bmatrix} x + y + 2z \\ x + 2y - z \\ x + 3y + z \end{bmatrix}$$

Setting corresponding entries equal to each other yields the following equivalent system:

$$\begin{aligned} x + y + 2z &= 1 \\ x + 2y - z &= -2 \\ x + 3y + z &= 5 \end{aligned} \tag{**}$$

For notational convenience, we have written the vectors in $\mathbf{R}^n$ as columns, because it is then easier to find the equivalent system of linear equations. In fact, one can easily go from the vector equation (*) directly to the system (**).

Now we solve the equivalent system of linear equations by reducing the system to echelon form. This yields

$$\begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 2y - z &= 4 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x + y + 2z &= 1 \\ y - 3z &= -3 \\ 5z &= 10 \end{aligned}$$

Back-substitution yields the solution $x = -6$, $y = 3$, $z = 2$. Thus, $v = -6u_1 + 3u_2 + 2u_3$.

EXAMPLE 3.15

(a) Write the vector $v = (4, 9, 19)$ as a linear combination of

$$u_1 = (1, -2, 3), \qquad u_2 = (3, -7, 10), \qquad u_3 = (2, 1, 9).$$

Find the equivalent system of linear equations by writing $v = xu_1 + yu_2 + zu_3$, and reduce the system to an echelon form. We have

$$\begin{array}{r} x + 3y + 2z = 4 \\ -2x - 7y + z = 9 \\ 3x + 10y + 9z = 19 \end{array} \quad \text{or} \quad \begin{array}{r} x + 3y + 2z = 4 \\ -y + 5z = 17 \\ y + 3z = 7 \end{array} \quad \text{or} \quad \begin{array}{r} x + 3y + 2z = 4 \\ -y + 5z = 17 \\ 8z = 24 \end{array}$$

Back-substitution yields the solution $x = 4$, $y = -2$, $z = 3$. Thus, v is a linear combination of u_1, u_2, u_3. Specifically, $v = 4u_1 - 2u_2 + 3u_3$.

(b) Write the vector $v = (2, 3, -5)$ as a linear combination of

$$u_1 = (1, 2, -3), \qquad u_2 = (2, 3, -4), \qquad u_3 = (1, 3, -5)$$

Find the equivalent system of linear equations by writing $v = xu_1 + yu_2 + zu_3$, and reduce the system to an echelon form. We have

$$\begin{array}{r} x + 2y + z = 2 \\ 2x + 3y + 3z = 3 \\ -3x - 4y - 5z = -5 \end{array} \quad \text{or} \quad \begin{array}{r} x + 2y + z = 2 \\ -y + z = -1 \\ 2y - 2z = 1 \end{array} \quad \text{or} \quad \begin{array}{r} x + 2y + z = 2 \\ -5y + 5z = -1 \\ 0 = 3 \end{array}$$

The system has no solution. Thus, it is impossible to write v as a linear combination of u_1, u_2, u_3.

Linear Combinations of Orthogonal Vectors, Fourier Coefficients

Recall first (Section 1.4) that the dot (inner) product $u \cdot v$ of vectors $u = (a_1, \ldots, a_n)$ and $v = (b_1, \ldots, b_n)$ in $\mathbf{R}^n$ is defined by

$$u \cdot v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Furthermore, vectors u and v are said to be *orthogonal* if their dot product $u \cdot v = 0$.

Suppose that $u_1, u_2, \ldots, u_n$ in $\mathbf{R}^n$ are n nonzero pairwise orthogonal vectors. This means

$$\text{(i)} \quad u_i \cdot u_j = 0 \quad \text{for } i \neq j \qquad \text{and} \qquad \text{(ii)} \quad u_i \cdot u_i \neq 0 \quad \text{for each } i$$

Then, for any vector v in $\mathbf{R}^n$, there is an easy way to write v as a linear combination of $u_1, u_2, \ldots, u_n$, which is illustrated in the next example.

EXAMPLE 3.16 Consider the following three vectors in $\mathbf{R}^3$:

$$u_1 = (1, 1, 1), \qquad u_2 = (1, -3, 2), \qquad u_3 = (5, -1, -4)$$

These vectors are pairwise orthogonal; that is,

$$u_1 \cdot u_2 = 1 - 3 + 2 = 0, \qquad u_1 \cdot u_3 = 5 - 1 - 4 = 0, \qquad u_2 \cdot u_3 = 5 + 3 - 8 = 0$$

Suppose we want to write $v = (4, 14, -9)$ as a linear combination of u_1, u_2, u_3.

Method 1. Find the equivalent system of linear equations as in Example 3.14 and then solve, obtaining $v = 3u_1 - 4u_2 + u_3$.

Method 2. (This method uses the fact that the vectors u_1, u_2, u_3 are mutually orthogonal, and hence, the arithmetic is much simpler.) Set v as a linear combination of u_1, u_2, u_3 using unknown scalars x, y, z as follows:

$$(4, 14, -9) = x(1, 1, 1) + y(1, -3, 2) + z(5, -1, -4) \qquad (*)$$

Take the dot product of (*) with respect to u_1 to get

$$(4, 14, -9) \cdot (1, 1, 1) = x(1, 1, 1) \cdot (1, 1, 1) \qquad \text{or} \qquad 9 = 3x \qquad \text{or} \qquad x = 3$$

(The last two terms drop out, because u_1 is orthogonal to u_2 and to u_3.) Next take the dot product of (*) with respect to u_2 to obtain

$$(4, 14, -9) \cdot (1, -3, 2) = y(1, -3, 2) \cdot (1, -3, 2) \qquad \text{or} \qquad -56 = 14y \qquad \text{or} \qquad y = -4$$

Finally, take the dot product of (*) with respect to u_3 to get

$$(4, 14, -9) \cdot (5, -1, -4) = z(5, -1, -4) \cdot (5, -1, -4) \qquad \text{or} \qquad 42 = 42z \qquad \text{or} \qquad z = 1$$

Thus, $v = 3u_1 - 4u_2 + u_3$.

The procedure in Method 2 in Example 3.16 is valid in general. Namely,

THEOREM 3.12: Suppose $u_1, u_2, \ldots, u_n$ are nonzero mutually orthogonal vectors in $\mathbf{R}^n$. Then, for any vector v in $\mathbf{R}^n$,

$$v = \frac{v \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{v \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{v \cdot u_n}{u_n \cdot u_n} u_n$$

We emphasize that there must be n such orthogonal vectors u_i in $\mathbf{R}^n$ for the formula to be used. Note also that each $u_i \cdot u_i \neq 0$, because each u_i is a nonzero vector.

Remark: The following scalar k_i (appearing in Theorem 3.12) is called the *Fourier coefficient* of v with respect to u_i:

$$k_i = \frac{v \cdot u_i}{u_i \cdot u_i} = \frac{v \cdot u_i}{\|u_i\|^2}$$

It is analogous to a coefficient in the celebrated Fourier series of a function.

3.11 Homogeneous Systems of Linear Equations

A system of linear equations is said to be *homogeneous* if all the constant terms are zero. Thus, a homogeneous system has the form $AX = 0$. Clearly, such a system always has the zero vector $0 = (0, 0, \ldots, 0)$ as a solution, called the *zero* or *trivial* solution. Accordingly, we are usually interested in whether or not the system has a nonzero solution.

Because a homogeneous system $AX = 0$ has at least the zero solution, it can always be put in an echelon form, say

$$\begin{array}{r} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 + \cdots + a_{1n}x_n = 0 \\ a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n = 0 \end{array}$$

Here r denotes the number of equations in echelon form and n denotes the number of unknowns. Thus, the echelon system has $n - r$ free variables.

The question of nonzero solutions reduces to the following two cases:

(i) $r = n$. The system has only the zero solution.

(ii) $r < n$. The system has a nonzero solution.

Accordingly, if we begin with fewer equations than unknowns, then, in echelon form, $r < n$, and the system has a nonzero solution. This proves the following important result.

THEOREM 3.13: A homogeneous system $AX = 0$ with more unknowns than equations has a nonzero solution.

EXAMPLE 3.17 Determine whether or not each of the following homogeneous systems has a nonzero solution:

$$\begin{array}{c} x + y - z = 0 \\ 2x - 3y + z = 0 \\ x - 4y + 2z = 0 \\ \text{(a)} \end{array} \qquad \begin{array}{c} x + y - z = 0 \\ 2x + 4y - z = 0 \\ 3x + 2y + 2z = 0 \\ \text{(b)} \end{array} \qquad \begin{array}{c} x_1 + 2x_2 - 3x_3 + 4x_4 = 0 \\ 2x_1 - 3x_2 + 5x_3 - 7x_4 = 0 \\ 5x_1 + 6x_2 - 9x_3 + 8x_4 = 0 \\ \text{(c)} \end{array}$$

(a) Reduce the system to echelon form as follows:

$$\begin{array}{r} x + y - z = 0 \\ -5y + 3z = 0 \\ -5y + 3z = 0 \end{array} \qquad \text{and then} \qquad \begin{array}{r} x + y - z = 0 \\ -5y + 3z = 0 \end{array}$$

The system has a nonzero solution, because there are only two equations in the three unknowns in echelon form. Here z is a free variable. Let us, say, set $z = 5$. Then, by back-substitution, $y = 3$ and $x = 2$. Thus, the vector $u = (2, 3, 5)$ is a particular nonzero solution.

(b) Reduce the system to echelon form as follows:

$$\begin{array}{r} x + y - z = 0 \\ 2y + z = 0 \\ -y + 5z = 0 \end{array} \qquad \text{and then} \qquad \begin{array}{r} x + y - z = 0 \\ 2y + z = 0 \\ 11z = 0 \end{array}$$

In echelon form, there are three equations in three unknowns. Thus, the system has only the zero solution.

(c) The system must have a nonzero solution (Theorem 3.13), because there are four unknowns but only three equations. (Here we do not need to reduce the system to echelon form.)

Basis for the General Solution of a Homogeneous System

Let W denote the general solution of a homogeneous system $AX = 0$. A list of nonzero solution vectors $u_1, u_2, \ldots, u_s$ of the system is said to be a *basis* for W if each solution vector $w \in W$ can be expressed uniquely as a linear combination of the vectors $u_1, u_2, \ldots, u_s$; that is, there exist unique scalars $a_1, a_2, \ldots, a_s$ such that

$$w = a_1 u_1 + a_2 u_2 + \cdots + a_s u_s$$

The number s of such basis vectors is equal to the number of free variables. This number s is called the *dimension* of W, written as $\dim W = s$. When $W = \{0\}$—that is, the system has only the zero solution—we define $\dim W = 0$.

The following theorem, proved in Chapter 5, page 171, tells us how to find such a basis.

THEOREM 3.14: Let W be the general solution of a homogeneous system $AX = 0$, and suppose that the echelon form of the homogeneous system has s free variables. Let $u_1, u_2, \ldots, u_s$ be the solutions obtained by setting one of the free variables equal to 1 (or any nonzero constant) and the remaining free variables equal to 0. Then $\dim W = s$, and the vectors $u_1, u_2, \ldots, u_s$ form a basis of W.

We emphasize that the general solution W may have many bases, and that Theorem 3.12 only gives us one such basis.

EXAMPLE 3.18 Find the dimension and a basis for the general solution W of the homogeneous system

$$\begin{array}{r} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 = 0 \\ 2x_1 + 4x_2 - 5x_3 + x_4 - 6x_5 = 0 \\ 5x_1 + 10x_2 - 13x_3 + 4x_4 - 16x_5 = 0 \end{array}$$

First reduce the system to echelon form. Apply the following operations:

"Replace L_2 by $-2L_1 + L_2$" and "Replace L_3 by $-5L_1 + L_3$" and then "Replace L_3 by $-2L_2 + L_3$"

These operations yield

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0 \\ x_3 - 3x_4 + 2x_5 &= 0 \\ 2x_3 - 6x_4 + 4x_5 &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 + 2x_2 - 3x_3 + 2x_4 - 4x_5 &= 0 \\ x_3 - 3x_4 + 2x_5 &= 0 \end{aligned}$$

The system in echelon form has three free variables, x_2, x_4, x_5; hence, $\dim W = 3$. Three solution vectors that form a basis for W are obtained as follows:

(1) Set $x_2 = 1$, $x_4 = 0$, $x_5 = 0$. Back-substitution yields the solution $u_1 = (-2, 1, 0, 0, 0)$.
(2) Set $x_2 = 0$, $x_4 = 1$, $x_5 = 0$. Back-substitution yields the solution $u_2 = (7, 0, 3, 1, 0)$.
(3) Set $x_2 = 0$, $x_4 = 0$, $x_5 = 1$. Back-substitution yields the solution $u_3 = (-2, 0, -2, 0, 1)$.

The vectors $u_1 = (-2, 1, 0, 0, 0)$, $u_2 = (7, 0, 3, 1, 0)$, $u_3 = (-2, 0, -2, 0, 1)$ form a basis for W.

Remark: Any solution of the system in Example 3.18 can be written in the form

$$\begin{aligned} au_1 + bu_2 + cu_3 &= a(-2, 1, 0, 0, 0) + b(7, 0, 3, 1, 0) + c(-2, 0, -2, 0, 1) \\ &= (-2a + 7b - 2c, \quad a, \quad 3b - 2c, \quad b, \quad c) \end{aligned}$$

or

$$x_1 = -2a + 7b - 2c, \qquad x_2 = a, \qquad x_3 = 3b - 2c, \qquad x_4 = b, \qquad x_5 = c$$

where a, b, c are arbitrary constants. Observe that this representation is nothing more than the parametric form of the general solution under the choice of parameters $x_2 = a$, $x_4 = b$, $x_5 = c$.

Nonhomogeneous and Associated Homogeneous Systems

Let $AX = B$ be a nonhomogeneous system of linear equations. Then $AX = 0$ is called the *associated homogeneous system*. For example,

$$\begin{aligned} x + 2y - 4z &= 7 \\ 3x - 5y + 6z &= 8 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x + 2y - 4z &= 0 \\ 3x - 5y + 6z &= 0 \end{aligned}$$

show a nonhomogeneous system and its associated homogeneous system.

The relationship between the solution U of a nonhomogeneous system $AX = B$ and the solution W of its associated homogeneous system $AX = 0$ is contained in the following theorem.

THEOREM 3.15: Let v_0 be a particular solution of $AX = B$ and let W be the general solution of $AX = 0$. Then the following is the general solution of $AX = B$:

$$U = v_0 + W = \{v_0 + w : w \in W\}$$

That is, $U = v_0 + W$ is obtained by adding v_0 to each element in W. We note that this theorem has a geometrical interpretation in $\mathbf{R}^3$. Specifically, suppose W is a line through the origin O. Then, as pictured in Fig. 3-4, $U = v_0 + W$ is the line parallel to W obtained by adding v_0 to each element of W. Similarly, whenever W is a plane through the origin O, then $U = v_0 + W$ is a plane parallel to W.

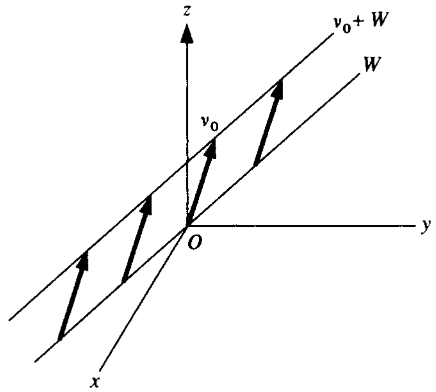


Figure 3-4

3.12 Elementary Matrices

Let e denote an elementary row operation and let $e(A)$ denote the results of applying the operation e to a matrix A. Now let E be the matrix obtained by applying e to the identity matrix I; that is,

$$E = e(I)$$

Then E is called the *elementary matrix* corresponding to the elementary row operation e. Note that E is always a square matrix.

EXAMPLE 3.19 Consider the following three elementary row operations:

(1) Interchange R_2 and R_3. (2) Replace R_2 by $-6R_2$. (3) Replace R_3 by $-4R_1 + R_3$.

The 3×3 elementary matrices corresponding to the above elementary row operations are as follows:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$$

The following theorem, proved in Problem 3.34, holds.

THEOREM 3.16: Let e be an elementary row operation and let E be the corresponding $m \times m$ elementary matrix. Then

$$e(A) = EA$$

where A is any $m \times n$ matrix.

In other words, the result of applying an elementary row operation e to a matrix A can be obtained by premultiplying A by the corresponding elementary matrix E.

Now suppose e' is the inverse of an elementary row operation e, and let E' and E be the corresponding matrices. We note (Problem 3.33) that E is invertible and E' is its inverse. This means, in particular, that any product

$$P = E_k \ldots E_2 E_1$$

of elementary matrices is invertible.

Applications of Elementary Matrices

Using Theorem 3.16, we are able to prove (Problem 3.35) the following important properties of matrices.

THEOREM 3.17: Let A be a square matrix. Then the following are equivalent:

(a) A is invertible (nonsingular).

(b) A is row equivalent to the identity matrix I.

(c) A is a product of elementary matrices.

Recall that square matrices A and B are inverses if $AB = BA = I$. The next theorem (proved in Problem 3.36) demonstrates that we need only show that one of the products is true, say $AB = I$, to prove that matrices are inverses.

THEOREM 3.18: Suppose $AB = I$. Then $BA = I$, and hence, $B = A^{-1}$.

Row equivalence can also be defined in terms of matrix multiplication. Specifically, we will prove (Problem 3.37) the following.

THEOREM 3.19: B is row equivalent to A if and only if there exists a nonsingular matrix P such that $B = PA$.

Application to Finding the Inverse of an $n \times n$ Matrix

The following algorithm finds the inverse of a matrix.

ALGORITHM 3.5: The input is a square matrix A. The output is the inverse of A or that the inverse does not exist.

Step 1. Form the $n \times 2n$ (block) matrix $M = [A, I]$, where A is the left half of M and the identity matrix I is the right half of M.

Step 2. Row reduce M to echelon form. If the process generates a zero row in the A half of M, then

STOP

A has no inverse. (Otherwise A is in triangular form.)

Step 3. Further row reduce M to its row canonical form

$$M \sim [I, B]$$

where the identity matrix I has replaced A in the left half of M.

Step 4. Set $A^{-1} = B$, the matrix that is now in the right half of M.

The justification for the above algorithm is as follows. Suppose A is invertible and, say, the sequence of elementary row operations $e_1, e_2, \dots, e_q$ applied to $M = [A, I]$ reduces the left half of M, which is A, to the identity matrix I. Let E_i be the elementary matrix corresponding to the operation e_i. Then, by applying Theorem 3.16. we get

$$E_q \dots E_2 E_1 A = I \qquad \text{or} \qquad (E_q \dots E_2 E_1 I)A = I, \qquad \text{so} \qquad A^{-1} = E_q \dots E_2 E_1 I$$

That is, A^{-1} can be obtained by applying the elementary row operations $e_1, e_2, \dots, e_q$ to the identity matrix I, which appears in the right half of M. Thus, $B = A^{-1}$, as claimed.

EXAMPLE 3.20

Find the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{bmatrix}$.

First form the (block) matrix $M = [A, I]$ and row reduce M to an echelon form:

$$M = \left[\begin{array}{rrr|rrr} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & -1 & 3 & 0 & 1 & 0 \\ 4 & 1 & 8 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -4 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -6 & 1 & 1 \end{array}\right]$$

In echelon form, the left half of M is in triangular form; hence, A has an inverse. Next we further row reduce M to its row canonical form:

$$M \sim \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & -1 & 0 & 4 & 0 & -1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right] \sim \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -11 & 2 & 2 \\ 0 & 1 & 0 & -4 & 0 & 1 \\ 0 & 0 & 1 & 6 & -1 & -1 \end{array}\right]$$

The identity matrix is now in the left half of the final matrix; hence, the right half is A^{-1}. In other words,

$$A^{-1} = \begin{bmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{bmatrix}$$

Elementary Column Operations

Now let A be a matrix with columns $C_1, C_2, \ldots, C_n$. The following operations on A, analogous to the elementary row operations, are called *elementary column operations*:

[F_1] (Column Interchange): Interchange columns C_i and C_j.
[F_2] (Column Scaling): Replace C_i by kC_i (where $k \neq 0$).
[F_3] (Column Addition): Replace C_j by $kC_i + C_j$.

We may indicate each of the column operations by writing, respectively,

$$(1)\ C_i \leftrightarrow C_j, \qquad (2)\ kC_i \rightarrow C_i, \qquad (3)\ (kC_i + C_j) \rightarrow C_j$$

Moreover, each column operation has an inverse operation of the same type, just like the corresponding row operation.

Now let f denote an elementary column operation, and let F be the matrix obtained by applying f to the identity matrix I; that is,

$$F = f(I)$$

Then F is called the *elementary matrix* corresponding to the elementary column operation f. Note that F is always a square matrix.

EXAMPLE 3.21
Consider the following elementary column operations:

(1) Interchange C_1 and C_3; (2) Replace C_3 by $-2C_3$; (3) Replace C_3 by $-3C_2 + C_3$

The corresponding three 3×3 elementary matrices are as follows:

$$F_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad F_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \qquad F_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

The following theorem is analogous to Theorem 3.16 for the elementary row operations.

THEOREM 3.20: For any matrix A, $f(A) = AF$.

That is, the result of applying an elementary column operation f on a matrix A can be obtained by postmultiplying A by the corresponding elementary matrix F.

Matrix Equivalence

A matrix B is *equivalent* to a matrix A if B can be obtained from A by a sequence of row and column operations. Alternatively, B is equivalent to A, if there exist nonsingular matrices P and Q such that $B = PAQ$. Just like row equivalence, equivalence of matrices is an equivalence relation.

The main result of this subsection (proved in Problem 3.38) is as follows.

THEOREM 3.21: Every $m \times n$ matrix A is equivalent to a unique block matrix of the form

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

where I_r is the r-square identity matrix.

The following definition applies.

DEFINITION: The nonnegative integer r in Theorem 3.18 is called the *rank* of A, written $\operatorname{rank}(A)$.

Note that this definition agrees with the previous definition of the rank of a matrix.

3.13 *LU* DECOMPOSITION

Suppose A is a nonsingular matrix that can be brought into (upper) triangular form U using only row-addition operations; that is, suppose A can be triangularized by the following algorithm, which we write using computer notation.

ALGORITHM 3.6: The input is a matrix A and the output is a triangular matrix U.

Step 1. Repeat for $i = 1, 2, \ldots, n-1$:

Step 2. Repeat for $j = i+1,\ i+2, \ldots, n$

(a) Set $m_{ij} := -a_{ij}/a_{ii}$.
(b) Set $R_j := m_{ij}R_i + R_j$

[End of Step 2 inner loop.]

[End of Step 1 outer loop.]

The numbers m_{ij} are called *multipliers*. Sometimes we keep track of these multipliers by means of the following lower triangular matrix L:

$$L = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ -m_{21} & 1 & 0 & \ldots & 0 & 0 \\ -m_{31} & -m_{32} & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -m_{n1} & -m_{n2} & -m_{n3} & \ldots & -m_{n,n-1} & 1 \end{bmatrix}$$

That is, L has 1's on the diagonal, 0's above the diagonal, and the negative of the multiplier m_{ij} as its ij-entry below the diagonal.

The above matrix L and the triangular matrix U obtained in Algorithm 3.6 give us the classical LU factorization of such a matrix A. Namely,

THEOREM 3.22: Let A be a nonsingular matrix that can be brought into triangular form U using only row-addition operations. Then $A = LU$, where L is the above lower triangular matrix with 1's on the diagonal, and U is an upper triangular matrix with no 0's on the diagonal.

EXAMPLE 3.22 Suppose $A = \begin{bmatrix} 1 & 2 & -3 \\ -3 & -4 & 13 \\ 2 & 1 & -5 \end{bmatrix}$. We note that A may be reduced to triangular form by the operations

"Replace R_2 by $3R_1 + R_2$"; "Replace R_3 by $-2R_1 + R_3$"; and then "Replace R_3 by $\frac{3}{2}R_2 + R_3$"

That is,

$$A \sim \begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{bmatrix}$$

This gives us the classical factorization $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -\frac{3}{2} & 1 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 2 & 4 \\ 0 & 0 & 7 \end{bmatrix}$$

We emphasize:

(1) The entries $-3, 2, -\frac{3}{2}$ in L are the negatives of the multipliers in the above elementary row operations.

(2) U is the triangular form of A.

Application to Systems of Linear Equations

Consider a computer algorithm M. Let $C(n)$ denote the running time of the algorithm as a function of the size n of the input data. [The function $C(n)$ is sometimes called the *time complexity* or simply the *complexity* of the algorithm M.] Frequently, $C(n)$ simply counts the number of multiplications and divisions executed by M, but does not count the number of additions and subtractions because they take much less time to execute.

Now consider a square system of linear equations $AX = B$, where

$$A = [a_{ij}], \qquad X = [x_1, \ldots, x_n]^T, \qquad B = [b_1, \ldots, b_n]^T$$

and suppose A has an LU factorization. Then the system can be brought into triangular form (in order to apply back-substitution) by applying Algorithm 3.6 to the augmented matrix $M = [A, B]$ of the system. The time complexity of Algorithm 3.6 and back-substitution are, respectively,

$$C(n) \approx \tfrac{1}{2}n^3 \qquad \text{and} \qquad C(n) \approx \tfrac{1}{2}n^2$$

where n is the number of equations.

On the other hand, suppose we already have the factorization $A = LU$. Then, to triangularize the system, we need only apply the row operations in the algorithm (retained by the matrix L) to the column vector B. In this case, the time complexity is

$$C(n) \approx \tfrac{1}{2}n^2$$

Of course, to obtain the factorization $A = LU$ requires the original algorithm where $C(n) \approx \frac{1}{2}n^3$. Thus, nothing may be gained by first finding the LU factorization when a single system is involved. However, there are situations, illustrated below, where the LU factorization is useful.

Suppose, for a given matrix A, we need to solve the system

$$AX = B$$

repeatedly for a sequence of different constant vectors, say $B_1, B_2, \ldots, B_k$. Also, suppose some of the B_i depend upon the solution of the system obtained while using preceding vectors B_j. In such a case, it is more efficient to first find the LU factorization of A, and then to use this factorization to solve the system for each new B.

EXAMPLE 3.23 Consider the following system of linear equations:

$$\begin{aligned} x + 2y + z &= k_1 \\ 2x + 3y + 3z &= k_2 \\ -3x + 10y + 2z &= k_3 \end{aligned} \quad \text{or} \quad AX = B, \quad \text{where} \quad A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ -3 & 10 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix}$$

Suppose we want to solve the system three times where B is equal, say, to B_1, B_2, B_3. Furthermore, suppose $B_1 = [1, 1, 1]^T$, and suppose

$$B_{j+1} = B_j + X_j \qquad (\text{for } j = 1, 2)$$

where X_j is the solution of $AX = B_j$. Here it is more efficient to first obtain the LU factorization of A and then use the LU factorization to solve the system for each of the B's. (This is done in Problem 3.42.)

SOLVED PROBLEMS

Linear Equations, Solutions, 2 × 2 Systems

3.1. Determine whether each of the following equations is linear:

(a) $5x + 7y - 8yz = 16$, (b) $x + \pi y + ez = \log 5$, (c) $3x + ky - 8z = 16$

(a) No, because the product yz of two unknowns is of second degree.

(b) Yes, because π, e, and $\log 5$ are constants.

(c) As it stands, there are four unknowns: x, y, z, k. Because of the term ky it is not a linear equation. However, assuming k is a constant, the equation is linear in the unknowns x, y, z.

3.2. Determine whether the following vectors are solutions of $x_1 + 2x_2 - 4x_3 + 3x_4 = 15$: (a) $u = (3, 2, 1, 4)$ and (b) $v = (1, 2, 4, 5)$.

(a) Substitute to obtain $3 + 2(2) - 4(1) + 3(4) = 15$, or $15 = 15$; yes, it is a solution.

(b) Substitute to obtain $1 + 2(2) - 4(4) + 3(5) = 15$, or $4 = 15$; no, it is not a solution.

3.3. Solve (a) $ex = \pi$, (b) $3x - 4 - x = 2x + 3$, (c) $7 + 2x - 4 = 3x + 3 - x$

(a) Because $e \neq 0$, multiply by $1/e$ to obtain $x = \pi/e$.

(b) Rewrite in standard form, obtaining $0x = 7$. The equation has no solution.

(c) Rewrite in standard form, obtaining $0x = 0$. Every scalar k is a solution.

3.4. Prove Theorem 3.4: Consider the equation $ax = b$.

(i) If $a \neq 0$, then $x = b/a$ is a unique solution of $ax = b$.

(ii) If $a = 0$ but $b \neq 0$, then $ax = b$ has no solution.

(iii) If $a = 0$ and $b = 0$, then every scalar k is a solution of $ax = b$.

Suppose $a \neq 0$. Then the scalar b/a exists. Substituting b/a in $ax = b$ yields $a(b/a) = b$, or $b = b$; hence, b/a is a solution. On the other hand, suppose x_0 is a solution to $ax = b$, so that $ax_0 = b$. Multiplying both sides by $1/a$ yields $x_0 = b/a$. Hence, b/a is the unique solution of $ax = b$. Thus, (i) is proved.

On the other hand, suppose $a = 0$. Then, for any scalar k, we have $ak = 0k = 0$. If $b \neq 0$, then $ak \neq b$. Accordingly, k is not a solution of $ax = b$, and so (ii) is proved. If $b = 0$, then $ak = b$. That is, any scalar k is a solution of $ax = b$, and so (iii) is proved.

3.5. Solve each of the following systems:

(a) $\begin{aligned} 2x - 5y &= 11 \\ 3x + 4y &= 5 \end{aligned}$ (b) $\begin{aligned} 2x - 3y &= 8 \\ -6x + 9y &= 6 \end{aligned}$ (c) $\begin{aligned} 2x - 3y &= 8 \\ -4x + 6y &= -16 \end{aligned}$

(a) Eliminate x from the equations by forming the new equation $L = -3L_1 + 2L_2$. This yields the equation

$$23y = -23, \qquad \text{and so} \qquad y = -1$$

Substitute $y = -1$ in one of the original equations, say L_1, to get

$$2x - 5(-1) = 11 \qquad \text{or} \qquad 2x + 5 = 11 \qquad \text{or} \qquad 2x = 6 \qquad \text{or} \qquad x = 3$$

Thus, $x = 3$, $y = -1$ or the pair $u = (3, -1)$ is the unique solution of the system.

(b) Eliminate x from the equations by forming the new equation $L = 3L_1 + L_2$. This yields the equation

$$0x + 0y = 30$$

This is a degenerate equation with a nonzero constant; hence, this equation and the system have no solution. (Geometrically, the lines corresponding to the equations are parallel.)

(c) Eliminate x from the equations by forming the new equation $L = 2L_1 + L_2$. This yields the equation

$$0x + 0y = 0$$

This is a degenerate equation where the constant term is also zero. Thus, the system has an infinite number of solutions, which correspond to the solution of either equation. (Geometrically, the lines corresponding to the equations coincide.)

To find the general solution, set $y = a$ and substitute in L_1 to obtain

$$2x - 3a = 8 \qquad \text{or} \qquad 2x = 3a + 8 \qquad \text{or} \qquad x = \tfrac{3}{2}a + 4$$

Thus, the general solution is

$$x = \tfrac{3}{2}a + 4, \quad y = a \qquad \text{or} \qquad u = \left(\tfrac{3}{2}a + 4,\ a\right)$$

where a is any scalar.

3.6. Consider the system

$$\begin{aligned} x + ay &= 4 \\ ax + 9y &= b \end{aligned}$$

(a) For which values of a does the system have a unique solution?

(b) Find those pairs of values (a, b) for which the system has more than one solution.

(a) Eliminate x from the equations by forming the new equation $L = -aL_1 + L_2$. This yields the equation

$$(9 - a^2)y = b - 4a \tag{1}$$

The system has a unique solution if and only if the coefficient of y in (1) is not zero—that is, if $9 - a^2 \neq 0$ or if $a \neq \pm 3$.

(b) The system has more than one solution if both sides of (1) are zero. The left-hand side is zero when $a = \pm 3$. When $a = 3$, the right-hand side is zero when $b - 12 = 0$ or $b = 12$. When $a = -3$, the right-hand side is zero when $b + 12 - 0$ or $b = -12$. Thus, $(3, 12)$ and $(-3, -12)$ are the pairs for which the system has more than one solution.

Systems in Triangular and Echelon Form

3.7. Determine the pivot and free variables in each of the following systems:

$$\begin{aligned} 2x_1 - 3x_2 - 6x_3 - 5x_4 + 2x_5 &= 7 \\ x_3 + 3x_4 - 7x_5 &= 6 \\ x_4 - 2x_5 &= 1 \end{aligned} \qquad \begin{aligned} 2x - 6y + 7z &= 1 \\ 4y + 3z &= 8 \\ 2z &= 4 \end{aligned} \qquad \begin{aligned} x + 2y - 3z &= 2 \\ 2x + 3y + z &= 4 \\ 3x + 4y + 5z &= 8 \end{aligned}$$

(a) (b) (c)

(a) In echelon form, the leading unknowns are the pivot variables, and the others are the free variables. Here x_1, x_3, x_4 are the pivot variables, and x_2 and x_5 are the free variables.

(b) The leading unknowns are x, y, z, so they are the pivot variables. There are no free variables (as in any triangular system).

(c) The notion of pivot and free variables applies only to a system in echelon form.

3.8. Solve the triangular system in Problem 3.7(b).

Because it is a triangular system, solve by back-substitution.

(i) The last equation gives $z = 2$.

(ii) Substitute $z = 2$ in the second equation to get $4y + 6 = 8$ or $y = \frac{1}{2}$.

(iii) Substitute $z = 2$ and $y = \frac{1}{2}$ in the first equation to get

$$2x - 6\left(\frac{1}{2}\right) + 7(2) = 1 \qquad \text{or} \qquad 2x + 11 = 1 \qquad \text{or} \qquad x = -5$$

Thus, $x = -5$, $y = \frac{1}{2}$, $z = 2$ or $u = (-5, \frac{1}{2}, 2)$ is the unique solution to the system.

3.9. Solve the echelon system in Problem 3.7(a).

Assign parameters to the free variables, say $x_2 = a$ and $x_5 = b$, and solve for the pivot variables by back-substitution.

(i) Substitute $x_5 = b$ in the last equation to get $x_4 - 2b = 1$ or $x_4 = 2b + 1$.

(ii) Substitute $x_5 = b$ and $x_4 = 2b + 1$ in the second equation to get

$$x_3 + 3(2b + 1) - 7b = 6 \qquad \text{or} \qquad x_3 - b + 3 = 6 \qquad \text{or} \qquad x_3 = b + 3$$

(iii) Substitute $x_5 = b,\ x_4 = 2b + 1,\ x_3 = b + 3,\ x_2 = a$ in the first equation to get

$$2x_1 - 3a - 6(b + 3) - 5(2b + 1) + 2b = 7 \qquad \text{or} \qquad 2x_1 - 3a - 14b - 23 = 7$$
$$\text{or} \qquad x_1 = \tfrac{3}{2}a + 7b + 15$$

Thus,

$$x_1 = \frac{3}{2}a + 7b + 15, \qquad x_2 = a, \qquad x_3 = b + 3, \qquad x_4 = 2b + 1, \qquad x_5 = b$$
$$\text{or} \qquad u = \left(\frac{3}{2}a + 7b + 15, \quad a, \quad b + 3, \quad 2b + 1, \quad b\right)$$

is the parametric form of the general solution.

Alternatively, solving for the pivot variable x_1, x_3, x_4 in terms of the free variables x_2 and x_5 yields the following free-variable form of the general solution:

$$x_1 = \frac{3}{2}x_2 + 7x_5 + 15, \qquad x_3 = x_5 + 3, \qquad x_4 = 2x_5 + 1$$

3.10. Prove Theorem 3.6. Consider the system (3.4) of linear equations in echelon form with r equations and n unknowns.

(i) If $r = n$, then the system has a unique solution.

(ii) If $r < n$, then we can arbitrarily assign values to the $n - r$ free variable and solve uniquely for the r pivot variables, obtaining a solution of the system.

(i) Suppose $r = n$. Then we have a square system $AX = B$ where the matrix A of coefficients is (upper) triangular with nonzero diagonal elements. Thus, A is invertible. By Theorem 3.10, the system has a unique solution.

(ii) Assigning values to the $n - r$ free variables yields a triangular system in the pivot variables, which, by (i), has a unique solution.

Gaussian Elimination

3.11. Solve each of the following systems:

$$\begin{array}{rcl} x+2y-4z &=& -4 \\ 2x+5y-9z &=& -10 \\ 3x-2y+3z &=& 11 \end{array} \qquad \begin{array}{rcl} x+2y-3z &=& -1 \\ -3x+y-2z &=& -7 \\ 5x+3y-4z &=& 2 \end{array} \qquad \begin{array}{rcl} x+2y-3z &=& 1 \\ 2x+5y-8z &=& 4 \\ 3x+8y-13z &=& 7 \end{array}$$

(a) (b) (c)

Reduce each system to triangular or echelon form using Gaussian elimination:

(a) Apply ''Replace L_2 by $-2L_1+L_2$'' and ''Replace L_3 by $-3L_1+L_3$'' to eliminate x from the second and third equations, and then apply ''Replace L_3 by $8L_2+L_3$'' to eliminate y from the third equation. These operations yield

$$\begin{array}{rcl} x+2y-4z &=& -4 \\ y-z &=& -2 \\ -8y+15z &=& 23 \end{array} \quad \text{and then} \quad \begin{array}{rcl} x+2y-4z &=& -4 \\ y-z &=& -2 \\ 7z &=& 7 \end{array}$$

The system is in triangular form. Solve by back-substitution to obtain the unique solution $u=(2,-1,1)$.

(b) Eliminate x from the second and third equations by the operations ''Replace L_2 by $3L_1+L_2$'' and ''Replace L_3 by $-5L_1+L_3$.'' This gives the equivalent system

$$\begin{array}{rcl} x+2y-3z &=& -1 \\ 7y-11z &=& -10 \\ -7y+11z &=& 7 \end{array}$$

The operation ''Replace L_3 by L_2+L_3'' yields the following degenerate equation with a nonzero constant:

$$0x+0y+0z=-3$$

This equation and hence the system have no solution.

(c) Eliminate x from the second and third equations by the operations ''Replace L_2 by $-2L_1+L_2$'' and ''Replace L_3 by $-3L_1+L_3$.'' This yields the new system

$$\begin{array}{rcl} x+2y-3z &=& 1 \\ y-2z &=& 2 \\ 2y-4z &=& 4 \end{array} \quad \text{or} \quad \begin{array}{rcl} x+2y-3z &=& 1 \\ y-2z &=& 2 \end{array}$$

(The third equation is deleted, because it is a multiple of the second equation.) The system is in echelon form with pivot variables x and y and free variable z.

To find the parametric form of the general solution, set $z=a$ and solve for x and y by back-substitution. Substitute $z=a$ in the second equation to get $y=2+2a$. Then substitute $z=a$ and $y=2+2a$ in the first equation to get

$$x+2(2+2a)-3a=1 \quad \text{or} \quad x+4+a=1 \quad \text{or} \quad x=-3-a$$

Thus, the general solution is

$$x=-3-a, \quad y=2+2a, \quad z=a \quad \text{or} \quad u=(-3-a, \quad 2+2a, \quad a)$$

where a is a parameter.

3.12. Solve each of the following systems:

$$\begin{array}{rcl} x_1-3x_2+2x_3-x_4+2x_5 &=& 2 \\ 3x_1-9x_2+7x_3-x_4+3x_5 &=& 7 \\ 2x_1-6x_2+7x_3+4x_4-5x_5 &=& 7 \end{array} \qquad \begin{array}{rcl} x_1+2x_2-3x_3+4x_4 &=& 2 \\ 2x_1+5x_2-2x_3+x_4 &=& 1 \\ 5x_1+12x_2-7x_3+6x_4 &=& 3 \end{array}$$

(a) (b)

Reduce each system to echelon form using Gaussian elimination:

(a) Apply ''Replace L_2 by $-3L_1 + L_2$'' and ''Replace L_3 by $-2L_1 + L_3$'' to eliminate x from the second and third equations. This yields

$$\begin{aligned} x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\ x_3 + 2x_4 - 3x_5 &= 1 \\ 3x_3 + 6x_4 - 9x_5 &= 3 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\ x_3 + 2x_4 - 3x_5 &= 1 \end{aligned}$$

(We delete L_3, because it is a multiple of L_2.) The system is in echelon form with pivot variables x_1 and x_3 and free variables x_2, x_4, x_5.

To find the parametric form of the general solution, set $x_2 = a$, $x_4 = b$, $x_5 = c$, where a, b, c are parameters. Back-substitution yields $x_3 = 1 - 2b + 3c$ and $x_1 = 3a + 5b - 8c$. The general solution is

$$x_1 = 3a + 5b - 8c, \quad x_2 = a, \quad x_3 = 1 - 2b + 3c, \quad x_4 = b, \quad x_5 = c$$

or, equivalently, $u = (3a + 5b - 8c, \ a, \ 1 - 2b + 3c, \ b, \ c)$.

(b) Eliminate x_1 from the second and third equations by the operations ''Replace L_2 by $-2L_1 + L_2$'' and ''Replace L_3 by $-5L_1 + L_3$.'' This yields the system

$$\begin{aligned} x_1 + 2x_2 - 3x_3 + 4x_4 &= 2 \\ x_2 + 4x_3 - 7x_4 &= -3 \\ 2x_2 + 8x_3 - 14x_4 &= -7 \end{aligned}$$

The operation ''Replace L_3 by $-2L_2 + L_3$'' yields the degenerate equation $0 = -1$. Thus, the system has no solution (even though the system has more unknowns than equations).

3.13. Solve using the condensed format:

$$\begin{aligned} 2y + 3z &= 3 \\ x + y + z &= 4 \\ 4x + 8y - 3z &= 35 \end{aligned}$$

The condensed format follows:

	Number	Equation	Operation
(2)	$(\cancel{1})$	$2y + 3z = 3$	$L_1 \leftrightarrow L_2$
(1)	$(\cancel{2})$	$x + y + z = 4$	$L_1 \leftrightarrow L_2$
	(3)	$4x + 8y - 3z = 35$	
	$(3')$	$4y - 7z = 19$	Replace L_3 by $-4L_1 + L_3$
	$(3'')$	$-13z = 13$	Replace L_3 by $-2L_2 + L_3$

Here (1), (2), and $(3'')$ form a triangular system. (We emphasize that the interchange of L_1 and L_2 is accomplished by simply renumbering L_1 and L_2 as above.)

Using back-substitution with the triangular system yields $z = -1$ from L_3, $y = 3$ from L_2, and $x = 2$ from L_1. Thus, the unique solution of the system is $x = 2$, $y = 3$, $z = -1$ or the triple $u = (2, 3, -1)$.

3.14. Consider the system

$$\begin{aligned} x + 2y + z &= 3 \\ ay + 5z &= 10 \\ 2x + 7y + az &= b \end{aligned}$$

(a) Find those values of a for which the system has a unique solution.

(b) Find those pairs of values (a, b) for which the system has more than one solution.

Reduce the system to echelon form. That is, eliminate x from the third equation by the operation ''Replace L_3 by $-2L_1 + L_3$'' and then eliminate y from the third equation by the operation

"Replace L_3 by $-3L_2 + aL_3$." This yields

$$\begin{aligned} x + 2y \quad + \quad z &= 3 \\ ay \quad + 5z &= 10 \\ 3y + (a-2)z &= b - 6 \end{aligned} \qquad \text{and then} \qquad \begin{aligned} x + 2y + \quad z &= 3 \\ ay + 5z &= 10 \\ (a^2 - 2a - 15)z &= ab - 6a - 30 \end{aligned}$$

Examine the last equation $(a^2 - 2a - 15)z = ab - 6a - 30$.

(a) The system has a unique solution if and only if the coefficient of z is not zero; that is, if

$$a^2 - 2a - 15 = (a-5)(a+3) \neq 0 \qquad \text{or} \qquad a \neq 5 \quad \text{and} \quad a \neq -3.$$

(b) The system has more than one solution if both sides are zero. The left-hand side is zero when $a = 5$ or $a = -3$. When $a = 5$, the right-hand side is zero when $5b - 60 = 0$, or $b = 12$. When $a = -3$, the right-hand side is zero when $-3b - 12 = 0$, or $b = -4$. Thus, $(5, 12)$ and $(-3, -4)$ are the pairs for which the system has more than one solution.

Echelon Matrices, Row Equivalence, Row Canonical Form

3.15. Row reduce each of the following matrices to echelon form:

$$\text{(a)} \quad A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{bmatrix}, \qquad \text{(b)} \quad B = \begin{bmatrix} -4 & 1 & -6 \\ 1 & 2 & -5 \\ 6 & 3 & -4 \end{bmatrix}$$

(a) Use $a_{11} = 1$ as a pivot to obtain 0's below a_{11}; that is, apply the row operations "Replace R_2 by $-2R_1 + R_2$" and "Replace R_3 by $-3R_1 + R_3$." Then use $a_{23} = 4$ as a pivot to obtain a 0 below a_{23}; that is, apply the row operation "Replace R_3 by $-5R_2 + 4R_3$." These operations yield

$$A \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

The matrix is now in echelon form.

(b) Hand calculations are usually simpler if the pivot element equals 1. Therefore, first interchange R_1 and R_2. Next apply the operations "Replace R_2 by $4R_1 + R_2$" and "Replace R_3 by $-6R_1 + R_3$"; and then apply the operation "Replace R_3 by $R_2 + R_3$." These operations yield

$$B \sim \begin{bmatrix} 1 & 2 & -5 \\ -4 & 1 & -6 \\ 6 & 3 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & -9 & 26 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -5 \\ 0 & 9 & -26 \\ 0 & 0 & 0 \end{bmatrix}$$

The matrix is now in echelon form.

3.16. Describe the *pivoting* row-reduction algorithm. Also describe the advantages, if any, of using this pivoting algorithm.

The row-reduction algorithm becomes a pivoting algorithm if the entry in column j of greatest absolute value is chosen as the pivot a_{1j_1} and if one uses the row operation

$$(-a_{ij_1}/a_{1j_1})R_1 + R_i \to R_i$$

The main advantage of the pivoting algorithm is that the above row operation involves division by the (current) pivot a_{1j_1}, and, on the computer, roundoff errors may be substantially reduced when one divides by a number as large in absolute value as possible.

3.17. Let $A = \begin{bmatrix} 2 & -2 & 2 & 1 \\ -3 & 6 & 0 & -1 \\ 1 & -7 & 10 & 2 \end{bmatrix}$. Reduce A to echelon form using the pivoting algorithm.

First interchange R_1 and R_2 so that -3 can be used as the pivot, and then apply the operations "Replace R_2 by $\frac{2}{3}R_1 + R_2$" and "Replace R_3 by $\frac{1}{3}R_1 + R_3$." These operations yield

$$A \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 2 & -2 & 2 & 1 \\ 1 & -7 & 10 & 2 \end{bmatrix} \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 0 & 2 & 2 & \frac{1}{3} \\ 0 & -5 & 10 & \frac{5}{3} \end{bmatrix}$$

Now interchange R_2 and R_3 so that -5 can be used as the pivot, and then apply the operation "Replace R_3 by $\frac{2}{5}R_2 + R_3$." We obtain

$$A \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 0 & -5 & 10 & \frac{5}{3} \\ 0 & 2 & 2 & \frac{1}{3} \end{bmatrix} \sim \begin{bmatrix} -3 & 6 & 0 & -1 \\ 0 & -5 & 10 & \frac{5}{3} \\ 0 & 0 & 6 & 1 \end{bmatrix}$$

The matrix has been brought to echelon form using partial pivoting.

3.18. Reduce each of the following matrices to row canonical form:

$$\text{(a)}\quad A = \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 4 & 4 & 1 & 10 & 13 \\ 8 & 8 & -1 & 26 & 23 \end{bmatrix}, \qquad \text{(b)}\quad B = \begin{bmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 7 \end{bmatrix}$$

(a) First reduce A to echelon form by applying the operations "Replace R_2 by $-2R_1 + R_2$" and "Replace R_3 by $-4R_1 + R_3$," and then applying the operation "Replace R_3 by $-R_2 + R_3$." These operations yield

$$A \sim \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 3 & 2 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 0 & 4 & 2 \end{bmatrix}$$

Now use back-substitution on the echelon matrix to obtain the row canonical form of A. Specifically, first multiply R_3 by $\frac{1}{4}$ to obtain the pivot $a_{34} = 1$, and then apply the operations "Replace R_2 by $2R_3 + R_2$" and "Replace R_1 by $-6R_3 + R_1$." These operations yield

$$A \sim \begin{bmatrix} 2 & 2 & -1 & 6 & 4 \\ 0 & 0 & 3 & -2 & 5 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 3 & 0 & 6 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

Now multiply R_2 by $\frac{1}{3}$, making the pivot $a_{23} = 1$, and then apply "Replace R_1 by $R_2 + R_1$," yielding

$$A \sim \begin{bmatrix} 2 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

Finally, multiply R_1 by $\frac{1}{2}$, so the pivot $a_{11} = 1$. Thus, we obtain the following row canonical form of A:

$$A \sim \begin{bmatrix} 1 & 1 & 0 & 0 & \frac{3}{2} \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

(b) Because B is in echelon form, use back-substitution to obtain

$$B \sim \begin{bmatrix} 5 & -9 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & -9 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & -9 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The last matrix, which is the identity matrix I, is the row canonical form of B. (This is expected, because B is invertible, and so its row canonical form must be I.)

3.19. Describe the Gauss–Jordan elimination algorithm, which also row reduces an arbitrary matrix A to its row canonical form.

The Gauss–Jordan algorithm is similar in some ways to the Gaussian elimination algorithm, except that here each pivot is used to place 0's both below and above the pivot, not just below the pivot, before working with the next pivot. Also, one variation of the algorithm first *normalizes* each row—that is, obtains a unit pivot—before it is used to produce 0's in the other rows, rather than normalizing the rows at the end of the algorithm.

3.20. Let $A = \begin{bmatrix} 1 & -2 & 3 & 1 & 2 \\ 1 & 1 & 4 & -1 & 3 \\ 2 & 5 & 9 & -2 & 8 \end{bmatrix}$. Use Gauss–Jordan to find the row canonical form of A.

Use $a_{11} = 1$ as a pivot to obtain 0's below a_{11} by applying the operations "Replace R_2 by $-R_1 + R_2$" and "Replace R_3 by $-2R_1 + R_3$." This yields

$$A \sim \begin{bmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 3 & 1 & -2 & 1 \\ 0 & 9 & 3 & -4 & 4 \end{bmatrix}$$

Multiply R_2 by $\frac{1}{3}$ to make the pivot $a_{22} = 1$, and then produce 0's below and above a_{22} by applying the operations "Replace R_3 by $-9R_2 + R_3$" and "Replace R_1 by $2R_2 + R_1$." These operations yield

$$A \sim \begin{bmatrix} 1 & -2 & 3 & 1 & 2 \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 9 & 3 & -4 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{11}{3} & -\frac{1}{3} & \frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 2 & 1 \end{bmatrix}$$

Finally, multiply R_3 by $\frac{1}{2}$ to make the pivot $a_{34} = 1$, and then produce 0's above a_{34} by applying the operations "Replace R_2 by $\frac{2}{3}R_3 + R_2$" and "Replace R_1 by $\frac{1}{3}R_3 + R_1$." These operations yield

$$A \sim \begin{bmatrix} 1 & 0 & \frac{11}{3} & -\frac{1}{3} & \frac{8}{3} \\ 0 & 1 & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{11}{3} & 0 & \frac{17}{6} \\ 0 & 1 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{bmatrix}$$

which is the row canonical form of A.

Systems of Linear Equations in Matrix Form

3.21. Find the augmented matrix M and the coefficient matrix A of the following system:

$$\begin{aligned} x + 2y - 3z &= 4 \\ 3y - 4z + 7x &= 5 \\ 6z + 8x - 9y &= 1 \end{aligned}$$

First align the unknowns in the system, and then use the aligned system to obtain M and A. We have

$$\begin{aligned} x + 2y - 3z &= 4 \\ 7x + 3y - 4z &= 5; \\ 8x - 9y + 6z &= 1 \end{aligned} \quad \text{then} \quad M = \begin{bmatrix} 1 & 2 & -3 & 4 \\ 7 & 3 & -4 & 5 \\ 8 & -9 & 6 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 2 & -3 \\ 7 & 3 & -4 \\ 8 & -9 & 6 \end{bmatrix}$$

3.22. Solve each of the following systems using its augmented matrix M:

$$\begin{array}{ccc} \begin{aligned} x + 2y - z &= 3 \\ x + 3y + z &= 5 \\ 3x + 8y + 4z &= 17 \end{aligned} & \begin{aligned} x - 2y + 4z &= 2 \\ 2x - 3y + 5z &= 3 \\ 3x - 4y + 6z &= 7 \end{aligned} & \begin{aligned} x + y + 3z &= 1 \\ 2x + 3y - z &= 3 \\ 5x + 7y + z &= 7 \end{aligned} \\ \text{(a)} & \text{(b)} & \text{(c)} \end{array}$$

(a) Reduce the augmented matrix M to echelon form as follows:

$$M = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 1 & 3 & 1 & 5 \\ 3 & 8 & 4 & 17 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 2 & 7 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 3 & 4 \end{bmatrix}$$

Now write down the corresponding triangular system

$$\begin{aligned} x + 2y - z &= 3 \\ y + 2z &= 2 \\ 3z &= 4 \end{aligned}$$

and solve by back-substitution to obtain the unique solution

$$x = \tfrac{17}{3},\ y = -\tfrac{2}{3},\ z = \tfrac{4}{3} \qquad \text{or} \qquad u = \left(\tfrac{17}{3}, -\tfrac{2}{3}, \tfrac{4}{3}\right)$$

Alternately, reduce the echelon form of M to row canonical form, obtaining

$$M \sim \begin{bmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & \frac{4}{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & \frac{13}{3} \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & \frac{17}{3} \\ 0 & 1 & 0 & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} \end{bmatrix}$$

This also corresponds to the above solution.

(b) First reduce the augmented matrix M to echelon form as follows:

$$M = \begin{bmatrix} 1 & -2 & 4 & 2 \\ 2 & -3 & 5 & 3 \\ 3 & -4 & 6 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 4 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 2 & -6 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 4 & 2 \\ 0 & 1 & -3 & -1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

The third row corresponds to the degenerate equation $0x + 0y + 0z = 3$, which has no solution. Thus, ''DO NOT CONTINUE.'' The original system also has no solution. (Note that the echelon form indicates whether or not the system has a solution.)

(c) Reduce the augmented matrix M to echelon form and then to row canonical form:

$$M = \begin{bmatrix} 1 & 1 & 3 & 1 \\ 2 & 3 & -1 & 3 \\ 5 & 7 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 3 & 1 \\ 0 & 1 & -7 & 1 \\ 0 & 2 & -14 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 10 & 0 \\ 0 & 1 & -7 & 1 \end{bmatrix}$$

(The third row of the second matrix is deleted, because it is a multiple of the second row and will result in a zero row.) Write down the system corresponding to the row canonical form of M and then transfer the free variables to the other side to obtain the free-variable form of the solution:

$$\begin{aligned} x + 10z &= 0 \\ y - 7z &= 1 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x &= -10z \\ y &= 1 + 7z \end{aligned}$$

Here z is the only free variable. The parametric solution, using $z = a$, is as follows:

$$x = -10a,\ y = 1 + 7a,\ z = a \qquad \text{or} \qquad u = (-10a,\ 1 + 7a,\ a)$$

3.23. Solve the following system using its augmented matrix M:

$$\begin{aligned} x_1 + 2x_2 - 3x_3 - 2x_4 + 4x_5 &= 1 \\ 2x_1 + 5x_2 - 8x_3 - x_4 + 6x_5 &= 4 \\ x_1 + 4x_2 - 7x_3 + 5x_4 + 2x_5 &= 8 \end{aligned}$$

Reduce the augmented matrix M to echelon form and then to row canonical form:

$$M = \begin{bmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 2 & 5 & -8 & -1 & 6 & 4 \\ 1 & 4 & -7 & 5 & 2 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 2 & -4 & 7 & -2 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -3 & -2 & 4 & 1 \\ 0 & 1 & -2 & 3 & -2 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -3 & 0 & 8 & 7 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 0 & 24 & 21 \\ 0 & 1 & -2 & 0 & -8 & -7 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}$$

Write down the system corresponding to the row canonical form of M and then transfer the free variables to the other side to obtain the free-variable form of the solution:

$$\begin{aligned} x_1 + x_3 + \ 24x_5 &= 21 \\ x_2 - 2x_3 - \ 8x_5 &= -7 \\ x_4 + 2x_5 &= 3 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 &= 21 - x_3 - 24x_5 \\ x_2 &= -7 + 2x_3 + 8x_5 \\ x_4 &= 3 - 2x_5 \end{aligned}$$

Here x_1, x_2, x_4 are the pivot variables and x_3 and x_5 are the free variables. Recall that the parametric form of the solution can be obtained from the free-variable form of the solution by simply setting the free variables equal to parameters, say $x_3 = a$, $x_5 = b$. This process yields

$$x_1 = 21 - a - 24b,\ \ x_2 = -7 + 2a + 8b,\ \ x_3 = a,\ \ x_4 = 3 - 2b,\ \ x_5 = b$$

or

$$u = (21 - a - 24b,\ \ -7 + 2a + 8b,\ \ a,\ \ 3 - 2b,\ \ b)$$

which is another form of the solution.

Linear Combinations, Homogeneous Systems

3.24. Write v as a linear combination of u_1, u_2, u_3, where

(a) $v = (3, 10, 7)$ and $u_1 = (1, 3, -2), u_2 = (1, 4, 2), u_3 = (2, 8, 1)$;

(b) $v = (2, 7, 10)$ and $u_1 = (1, 2, 3),\ u_2 = (1, 3, 5),\ u_3 = (1, 5, 9)$;

(c) $v = (1, 5, 4)$ and $u_1 = (1, 3, -2),\ u_2 = (2, 7, -1),\ u_3 = (1, 6, 7)$.

Find the equivalent system of linear equations by writing $v = xu_1 + yu_2 + zu_3$. Alternatively, use the augmented matrix M of the equivalent system, where $M = [u_1, u_2, u_3, v]$. (Here u_1, u_2, u_3, v are the columns of M.)

(a) The vector equation $v = xu_1 + yu_2 + zu_3$ for the given vectors is as follows:

$$\begin{bmatrix} 3 \\ 10 \\ 7 \end{bmatrix} = x\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + y\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + z\begin{bmatrix} 2 \\ 8 \\ 1 \end{bmatrix} = \begin{bmatrix} x + y + 2z \\ 3x + 4y + 8z \\ -2x + 2y + z \end{bmatrix}$$

Form the equivalent system of linear equations by setting corresponding entries equal to each other, and then reduce the system to echelon form:

$$\begin{aligned} x + \ \ y + 2z &= \ \ 3 \\ 3x + 4y + 8z &= 10 \\ -2x + 2y + \ \ z &= \ \ 7 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + \ \ y + 2z &= \ \ 3 \\ y + 2z &= \ \ 1 \\ 4y + 5z &= 13 \end{aligned} \qquad \text{or} \qquad \begin{aligned} x + y + 2z &= 3 \\ y + 2z &= 1 \\ -3z &= 9 \end{aligned}$$

The system is in triangular form. Back-substitution yields the unique solution $x = 2$, $y = 7$, $z = -3$. Thus, $v = 2u_1 + 7u_2 - 3u_3$.

Alternatively, form the augmented matrix $M = [u_1, u_2, u_3, v]$ of the equivalent system, and reduce M to echelon form:

$$M = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 3 & 4 & 8 & 10 \\ -2 & 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 4 & 5 & 13 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -3 & 9 \end{bmatrix}$$

The last matrix corresponds to a triangular system that has a unique solution. Back-substitution yields the solution $x = 2$, $y = 7$, $z = -3$. Thus, $v = 2u_1 + 7u_2 - 3u_3$.

(b) Form the augmented matrix $M = [u_1, u_2, u_3, v]$ of the equivalent system, and reduce M to the echelon form:

$$M = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 2 & 3 & 5 & 7 \\ 3 & 5 & 9 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 3 & 3 \\ 0 & 2 & 6 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 3 & 3 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

The third row corresponds to the degenerate equation $0x + 0y + 0z = -2$, which has no solution. Thus, the system also has no solution, and v cannot be written as a linear combination of u_1, u_2, u_3.

(c) Form the augmented matrix $M = [u_1, u_2, u_3, v]$ of the equivalent system, and reduce M to echelon form:

$$M = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 3 & 7 & 6 & 5 \\ -2 & -1 & 7 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 3 & 9 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The last matrix corresponds to the following system with free variable z:

$$\begin{aligned} x + 2y + z &= 1 \\ y + 3z &= 2 \end{aligned}$$

Thus, v can be written as a linear combination of u_1, u_2, u_3 in many ways. For example, let the free variable $z = 1$, and, by back-substitution, we get $y = -2$ and $x = 2$. Thus, $v = 2u_1 - 2u_2 + u_3$.

3.25. Let $u_1 = (1, 2, 4)$, $u_2 = (2, -3, 1)$, $u_3 = (2, 1, -1)$ in $\mathbf{R}^3$. Show that u_1, u_2, u_3 are orthogonal, and write v as a linear combination of u_1, u_2, u_3, where (a) $v = (7, 16, 6)$, (b) $v = (3, 5, 2)$.

Take the dot product of pairs of vectors to get

$$u_1 \cdot u_2 = 2 - 6 + 4 = 0, \quad u_1 \cdot u_3 = 2 + 2 - 4 = 0, \quad u_2 \cdot u_3 = 4 - 3 - 1 = 0$$

Thus, the three vectors in $\mathbf{R}^3$ are orthogonal, and hence Fourier coefficients can be used. That is, $v = xu_1 + yu_2 + zu_3$, where

$$x = \frac{v \cdot u_1}{u_1 \cdot u_1}, \qquad y = \frac{v \cdot u_2}{u_2 \cdot u_2}, \qquad z = \frac{v \cdot u_3}{u_3 \cdot u_3}$$

(a) We have

$$x = \frac{7 + 32 + 24}{1 + 4 + 16} = \frac{63}{21} = 3, \qquad y = \frac{14 - 48 + 6}{4 + 9 + 1} = \frac{-28}{14} = -2, \qquad z = \frac{14 + 16 - 6}{4 + 1 + 1} = \frac{24}{6} = 4$$

Thus, $v = 3u_1 - 2u_2 + 4u_3$.

(b) We have

$$x = \frac{3 + 10 + 8}{1 + 4 + 16} = \frac{21}{21} = 1, \qquad y = \frac{6 - 15 + 2}{4 + 9 + 1} = \frac{-7}{14} = -\frac{1}{2}, \qquad z = \frac{6 + 5 - 2}{4 + 1 + 1} = \frac{9}{6} = \frac{3}{2}$$

Thus, $v = u_1 - \frac{1}{2}u_2 + \frac{3}{2}u_3$.

3.26. Find the dimension and a basis for the general solution W of each of the following homogeneous systems:

$$\begin{aligned} 2x_1 + 4x_2 - 5x_3 + 3x_4 &= 0 \\ 3x_1 + 6x_2 - 7x_3 + 4x_4 &= 0 \\ 5x_1 + 10x_2 - 11x_3 + 6x_4 &= 0 \end{aligned} \qquad\qquad \begin{aligned} x - 2y - 3z &= 0 \\ 2x + y + 3z &= 0 \\ 3x - 4y - 2z &= 0 \end{aligned}$$

(a) (b)

(a) Reduce the system to echelon form using the operations "Replace L_2 by $-3L_1 + 2L_2$," "Replace L_3 by $-5L_1 + 2L_3$," and then "Replace L_3 by $-2L_2 + L_3$." These operations yield

$$\begin{aligned} 2x_1 + 4x_2 - 5x_3 + 3x_4 &= 0 \\ x_3 - x_4 &= 0 \\ 3x_3 - 3x_4 &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 2x_1 + 4x_2 - 5x_3 + 3x_4 &= 0 \\ x_3 - x_4 &= 0 \end{aligned}$$

The system in echelon form has two free variables, x_2 and x_4, so $\dim W = 2$. A basis $[u_1, u_2]$ for W may be obtained as follows:

(1) Set $x_2 = 1$, $x_4 = 0$. Back-substitution yields $x_3 = 0$, and then $x_1 = -2$. Thus, $u_1 = (-2, 1, 0, 0)$.

(2) Set $x_2 = 0$, $x_4 = 1$. Back-substitution yields $x_3 = 1$, and then $x_1 = 1$. Thus, $u_2 = (1, 0, 1, 1)$.

(b) Reduce the system to echelon form, obtaining

$$\begin{aligned} x - 2y - 3z &= 0 \\ 5y + 9z &= 0 \\ 2y + 7z &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x - 2y - 3z &= 0 \\ 5y + 9z &= 0 \\ 17z &= 0 \end{aligned}$$

There are no free variables (the system is in triangular form). Hence, $\dim W = 0$, and W has no basis. Specifically, W consists only of the zero solution; that is, $W = \{0\}$.

3.27. Find the dimension and a basis for the general solution W of the following homogeneous system using matrix notation:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 - 2x_4 + 4x_5 &= 0 \\ 2x_1 + 4x_2 + 8x_3 + x_4 + 9x_5 &= 0 \\ 3x_1 + 6x_2 + 13x_3 + 4x_4 + 14x_5 &= 0 \end{aligned}$$

Show how the basis gives the parametric form of the general solution of the system.

When a system is homogeneous, we represent the system by its coefficient matrix A rather than by its

augmented matrix M, because the last column of the augmented matrix M is a zero column, and it will remain a zero column during any row-reduction process.

Reduce the coefficient matrix A to echelon form, obtaining

$$A = \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 2 & 4 & 8 & 1 & 9 \\ 3 & 6 & 13 & 4 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 0 & 2 & 5 & 1 \\ 0 & 0 & 4 & 10 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 0 & 2 & 5 & 1 \end{bmatrix}$$

(The third row of the second matrix is deleted, because it is a multiple of the second row and will result in a zero row.) We can now proceed in one of two ways.

(a) Write down the corresponding homogeneous system in echelon form:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 - 2x_4 + 4x_5 &= 0 \\ 2x_3 + 5x_4 + x_5 &= 0 \end{aligned}$$

The system in echelon form has three free variables, x_2, x_4, x_5, so $\dim W = 3$. A basis $[u_1, u_2, u_3]$ for W may be obtained as follows:

(1) Set $x_2 = 1$, $x_4 = 0$, $x_5 = 0$. Back-substitution yields $x_3 = 0$, and then $x_1 = -2$. Thus,
$$u_1 = (-2, 1, 0, 0, 0).$$

(2) Set $x_2 = 0$, $x_4 = 1$, $x_5 = 0$. Back-substitution yields $x_3 = -\frac{5}{2}$, and then $x_1 = \frac{19}{2}$. Thus,
$$u_2 = (\tfrac{19}{2}, 0, -\tfrac{5}{2}, 1, 0).$$

(3) Set $x_2 = 0$, $x_4 = 0$, $x_5 = 1$. Back-substitution yields $x_3 = -\frac{1}{2}$, and then $x_1 = -\frac{5}{2}$. Thus,
$$u_3 = (-\tfrac{5}{2}, 0, -\tfrac{1}{2}, 0, 1).$$

[One could avoid fractions in the basis by choosing $x_4 = 2$ in (2) and $x_5 = 2$ in (3), which yields multiples of u_2 and u_3.] The parametric form of the general solution is obtained from the following linear combination of the basis vectors using parameters a, b, c:

$$au_1 + bu_2 + cu_3 = (-2a + \tfrac{19}{2}b - \tfrac{5}{2}c,\ \ a,\ \ -\tfrac{5}{2}b - \tfrac{1}{2}c,\ \ b,\ \ c)$$

(b) Reduce the echelon form of A to row canonical form:

$$A \sim \begin{bmatrix} 1 & 2 & 3 & -2 & 4 \\ 0 & 0 & 1 & \frac{5}{2} & \frac{1}{2} \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & -\frac{19}{2} & \frac{5}{2} \\ 0 & 0 & 1 & \frac{5}{2} & \frac{1}{2} \end{bmatrix}$$

Write down the corresponding free-variable solution:

$$x_1 = -2x_2 + \frac{19}{2}x_4 - \frac{5}{2}x_5$$
$$x_3 = -\frac{5}{2}x_4 - \frac{1}{2}x_5$$

Using these equations for the pivot variables x_1 and x_3, repeat the above process to obtain a basis $[u_1, u_2, u_3]$ for W. That is, set $x_2 = 1$, $x_4 = 0$, $x_5 = 0$ to get u_1; set $x_2 = 0$, $x_4 = 1$, $x_5 = 0$ to get u_2; and set $x_2 = 0$, $x_4 = 0$, $x_5 = 1$ to get u_3.

3.28. Prove Theorem 3.15. Let v_0 be a particular solution of $AX = B$, and let W be the general solution of $AX = 0$. Then $U = v_0 + W = \{v_0 + w : w \in W\}$ is the general solution of $AX = B$.

Let w be a solution of $AX = 0$. Then

$$A(v_0 + w) = Av_0 + Aw = B + 0 = B$$

Thus, the sum $v_0 + w$ is a solution of $AX = B$. On the other hand, suppose v is also a solution of $AX = B$. Then

$$A(v - v_0) = Av - Av_0 = B - B = 0$$

Therefore, $v - v_0$ belongs to W. Because $v = v_0 + (v - v_0)$, we find that any solution of $AX = B$ can be obtained by adding a solution of $AX = 0$ to a solution of $AX = B$. Thus, the theorem is proved.

Elementary Matrices, Applications

3.29. Let e_1, e_2, e_3 denote, respectively, the elementary row operations

"Interchange rows R_1 and R_2," "Replace R_3 by $7R_3$," "Replace R_2 by $-3R_1 + R_2$"

Find the corresponding three-square elementary matrices E_1, E_2, E_3. Apply each operation to the 3×3 identity matrix I_3 to obtain

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}, \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3.30. Consider the elementary row operations in Problem 3.29.

(a) Describe the inverse operations $e_1^{-1}, e_2^{-1}, e_3^{-1}$.

(b) Find the corresponding three-square elementary matrices E_1', E_2', E_3'.

(c) What is the relationship between the matrices E_1', E_2', E_3' and the matrices E_1, E_2, E_3?

(a) The inverses of e_1, e_2, e_3 are, respectively,

"Interchange rows R_1 and R_2," "Replace R_3 by $\frac{1}{7}R_3$," "Replace R_2 by $3R_1 + R_2$."

(b) Apply each inverse operation to the 3×3 identity matrix I_3 to obtain

$$E_1' = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad E_2' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{7} \end{bmatrix}, \qquad E_3' = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(c) The matrices E_1', E_2', E_3' are, respectively, the inverses of the matrices E_1, E_2, E_3.

3.31. Write each of the following matrices as a product of elementary matrices:

(a) $A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$, (c) $C = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 8 \\ -3 & -1 & 2 \end{bmatrix}$

The following three steps write a matrix M as a product of elementary matrices:

Step 1. Row reduce M to the identity matrix I, keeping track of the elementary row operations.

Step 2. Write down the inverse row operations.

Step 3. Write M as the product of the elementary matrices corresponding to the inverse operations. This gives the desired result.

If a zero row appears in Step 1, then M is not row equivalent to the identity matrix I, and M cannot be written as a product of elementary matrices.

(a) (1) We have

$$A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 \\ 0 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

where the row operations are, respectively,

"Replace R_2 by $2R_1 + R_2$," "Replace R_2 by $-\frac{1}{2}R_2$," "Replace R_1 by $3R_2 + R_1$"

(2) Inverse operations:

"Replace R_2 by $-2R_1 + R_2$," "Replace R_2 by $-2R_2$," "Replace R_1 by $-3R_2 + R_1$"

(3) $A = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$

(b) (1) We have

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

where the row operations are, respectively,

"Replace R_2 by $-4R_3 + R_2$," "Replace R_1 by $-3R_3 + R_1$," "Replace R_1 by $-2R_2 + R_1$"

(2) Inverse operations:

"Replace R_2 by $4R_3 + R_2$," "Replace R_1 by $3R_3 + R_1$," "Replace R_1 by $2R_2 + R_1$"

(3) $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) (1) First row reduce C to echelon form. We have

$$C = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 8 \\ -3 & -1 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 4 \\ 0 & 2 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

In echelon form, C has a zero row. "STOP." The matrix C cannot be row reduced to the identity matrix I, and C cannot be written as a product of elementary matrices. (We note, in particular, that C has no inverse.)

3.32. Find the inverse of (a) $A = \begin{bmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & -3 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{bmatrix}$.

(a) Form the matrix $M = [A, I]$ and row reduce M to echelon form:

$$M = \left[\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ -1 & -1 & 5 & 0 & 1 & 0 \\ 2 & 7 & -3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 3 & 5 & -2 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & -4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 2 & -5 & -3 & 1 \end{array}\right]$$

In echelon form, the left half of M is in triangular form; hence, A has an inverse. Further reduce M to row canonical form:

$$M \sim \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -9 & -6 & 2 \\ 0 & 1 & 0 & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -16 & -11 & 3 \\ 0 & 1 & 0 & \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{array}\right]$$

The final matrix has the form $[I, A^{-1}]$; that is, A^{-1} is the right half of the last matrix. Thus,

$$A^{-1} = \begin{bmatrix} -16 & -11 & 3 \\ \frac{7}{2} & \frac{5}{2} & -\frac{1}{2} \\ -\frac{5}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

(b) Form the matrix $M = [B, I]$ and row reduce M to echelon form:

$$M = \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 1 & 5 & -1 & 0 & 1 & 0 \\ 3 & 13 & -6 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 0 \\ 0 & 4 & 6 & -3 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 3 & -4 & 1 & 0 & 0 \\ 0 & 2 & 3 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & -2 & 1 \end{array}\right]$$

In echelon form, M has a zero row in its left half; that is, B is not row reducible to triangular form. Accordingly, B has no inverse.

3.33. Show that every elementary matrix E is invertible, and its inverse is an elementary matrix.

Let E be the elementary matrix corresponding to the elementary operation e; that is, $e(I) = E$. Let e' be the inverse operation of e and let E' be the corresponding elementary matrix; that is, $e'(I) = E'$. Then

$$I = e'(e(I)) = e'(E) = E'E \qquad \text{and} \qquad I = e(e'(I)) = e(E') = EE'$$

Therefore, E' is the inverse of E.

3.34. Prove Theorem 3.16: Let e be an elementary row operation and let E be the corresponding m-square elementary matrix; that is, $E = e(I)$. Then $e(A) = EA$, where A is any $m \times n$ matrix.

Let R_i be the row i of A; we denote this by writing $A = [R_1, \ldots, R_m]$. If B is a matrix for which AB is defined then $AB = [R_1 B, \ldots, R_m B]$. We also let

$$e_i = (0, \ldots, 0, \hat{1}, 0, \ldots, 0), \qquad \hat{} = i$$

Here $\hat{} = i$ means 1 is the ith entry. One can show (Problem 2.45) that $e_i A = R_i$. We also note that $I = [e_1, e_2, \ldots, e_m]$ is the identity matrix.

(i) Let e be the elementary row operation "Interchange rows R_i and R_j." Then, for $\hat{} = i$ and $\hat{\hat{}} = j$,

$$E = e(I) = [e_1, \ldots, \widehat{e_j}, \ldots, \widehat{\widehat{e_i}}, \ldots, e_m]$$

and

$$e(A) = [R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m]$$

Thus,

$$EA = [e_1 A, \ldots, \widehat{e_j A}, \ldots, \widehat{\widehat{e_i A}}, \ldots, e_m A] = [R_1, \ldots, \widehat{R_j}, \ldots, \widehat{\widehat{R_i}}, \ldots, R_m] = e(A)$$

(ii) Let e be the elementary row operation "Replace R_i by kR_i $(k \neq 0)$." Then, for $\hat{} = i$,

$$E = e(I) = [e_1, \ldots, \widehat{ke_i}, \ldots, e_m]$$

and

$$e(A) = [R_1, \ldots, \widehat{kR_i}, \ldots, R_m]$$

Thus,

$$EA = [e_1 A, \ldots, \widehat{ke_i A}, \ldots, e_m A] = [R_1, \ldots, \widehat{kR_i}, \ldots, R_m] = e(A)$$

(iii) Let e be the elementary row operation "Replace R_i by $kR_j + R_i$." Then, for $\hat{} = i$,

$$E = e(I) = [e_1, \ \ldots, \ \widehat{ke_j + e_i}, \ \ldots, \ e_m]$$

and

$$e(A) = [R_1, \ \ldots, \ \widehat{kR_j + R_i}, \ \ldots, \ R_m]$$

Using $(ke_j + e_i)A = k(e_j A) + e_i A = kR_j + R_i$, we have

$$\begin{aligned} EA &= [e_1 A, \quad \ldots, \quad (ke_j + e_i)A, \quad \ldots, \quad e_m A] \\ &= [R_1, \quad \ldots, \quad \widehat{kR_j + R_i}, \quad \ldots, \quad R_m] = e(A) \end{aligned}$$

3.35. Prove Theorem 3.17: Let A be a square matrix. Then the following are equivalent:

(a) A is invertible (nonsingular).

(b) A is row equivalent to the identity matrix I.

(c) A is a product of elementary matrices.

Suppose A is invertible and suppose A is row equivalent to matrix B in row canonical form. Then there exist elementary matrices $E_1, E_2, \ldots, E_s$ such that $E_s \ldots E_2 E_1 A = B$. Because A is invertible and each elementary matrix is invertible, B is also invertible. But if $B \neq I$, then B has a zero row; whence B is not invertible. Thus, $B = I$, and (a) implies (b).

If (b) holds, then there exist elementary matrices $E_1, E_2, \ldots, E_s$ such that $E_s \ldots E_2 E_1 A = I$. Hence, $A = (E_s \ldots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \ldots, E_s^{-1}$. But the E_i^{-1} are also elementary matrices. Thus (b) implies (c).

If (c) holds, then $A = E_1 E_2 \ldots E_s$. The E_i are invertible matrices; hence, their product A is also invertible. Thus, (c) implies (a). Accordingly, the theorem is proved.

3.36. Prove Theorem 3.18: If $AB = I$, then $BA = I$, and hence $B = A^{-1}$.

Suppose A is not invertible. Then A is not row equivalent to the identity matrix I, and so A is row equivalent to a matrix with a zero row. In other words, there exist elementary matrices $E_1, \ldots, E_s$ such that $E_s \ldots E_2 E_1 A$ has a zero row. Hence, $E_s \ldots E_2 E_1 AB = E_s \ldots E_2 E_1$, an invertible matrix, also has a zero row. But invertible matrices cannot have zero rows; hence A is invertible, with inverse A^{-1}. Then also,

$$B = IB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}I = A^{-1}$$

3.37. Prove Theorem 3.19: B is row equivalent to A (written $B \sim A$) if and only if there exists a nonsingular matrix P such that $B = PA$.

If $B \sim A$, then $B = e_s(\ldots(e_2(e_1(A)))\ldots) = E_s \ldots E_2 E_1 A = PA$ where $P = E_s \ldots E_2 E_1$ is nonsingular. Conversely, suppose $B = PA$, where P is nonsingular. By Theorem 3.17, P is a product of elementary matrices, and so B can be obtained from A by a sequence of elementary row operations; that is, $B \sim A$. Thus, the theorem is proved.

3.38. Prove Theorem 3.21: Every $m \times n$ matrix A is equivalent to a unique block matrix of the form $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where I_r is the $r \times r$ identity matrix.

The proof is constructive, in the form of an algorithm.

Step 1. Row reduce A to row canonical form, with leading nonzero entries $a_{1j_1}, a_{2j_2}, \ldots, a_{rj_r}$.

Step 2. Interchange C_1 and C_{1j_1}, interchange C_2 and $C_{2j_2}, \ldots$, and interchange C_r and C_{jr}. This gives a matrix in the form $\left[\begin{array}{c|c} I_r & B \\ \hline 0 & 0 \end{array}\right]$, with leading nonzero entries $a_{11}, a_{22}, \ldots, a_{rr}$.

Step 3. Use column operations, with the a_{ii} as pivots, to replace each entry in B with a zero; that is, for $i = 1, 2, \ldots, r$ and $j = r+1, r+2, \ldots, n$, apply the operation $-b_{ij}C_i + C_j \to C_j$.

The final matrix has the desired form $\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$.

Lu Factorization

3.39. Find the LU factorization of (a) $A = \begin{bmatrix} 1 & -3 & 5 \\ 2 & -4 & 7 \\ -1 & -2 & 1 \end{bmatrix}$, (b) $B = \begin{bmatrix} 1 & 4 & -3 \\ 2 & 8 & 1 \\ -5 & -9 & 7 \end{bmatrix}$.

(a) Reduce A to triangular form by the following operations:

"Replace R_2 by $-2R_1 + R_2$," "Replace R_3 by $R_1 + R_3$," and then
"Replace R_3 by $\frac{5}{2}R_2 + R_3$"

These operations yield the following, where the triangular form is U:

$$A \sim \begin{bmatrix} 1 & -3 & 5 \\ 0 & 2 & -3 \\ 0 & -5 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 5 \\ 0 & 2 & -3 \\ 0 & 0 & -\frac{3}{2} \end{bmatrix} = U \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -\frac{5}{2} & 1 \end{bmatrix}$$

The entries $2, -1, -\frac{5}{2}$ in L are the negatives of the multipliers $-2, 1, \frac{5}{2}$ in the above row operations. (As a check, multiply L and U to verify $A = LU$.)

(b) Reduce B to triangular form by first applying the operations "Replace R_2 by $-2R_1 + R_2$" and "Replace R_3 by $5R_1 + R_3$." These operations yield

$$B \sim \begin{bmatrix} 1 & 4 & -3 \\ 0 & 0 & 7 \\ 0 & 11 & -8 \end{bmatrix}.$$

Observe that the second diagonal entry is 0. Thus, B cannot be brought into triangular form without row interchange operations. Accordingly, B is not LU-factorable. (There does exist a PLU factorization of such a matrix B, where P is a permutation matrix, but such a factorization lies beyond the scope of this text.)

3.40. Find the LDU factorization of the matrix A in Problem 3.39.

The $A = LDU$ factorization refers to the situation where L is a lower triangular matrix with 1's on the diagonal (as in the LU factorization of A), D is a diagonal matrix, and U is an upper triangular matrix with 1's on the diagonal. Thus, simply factor out the diagonal entries in the matrix U in the above LU factorization of A to obtain D and L. That is,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -\frac{5}{2} & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -\frac{3}{2} \end{bmatrix}, \qquad U = \begin{bmatrix} 1 & -3 & 5 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$$

3.41. Find the LU factorization of the matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 3 \\ -3 & -10 & 2 \end{bmatrix}$.

Reduce A to triangular form by the following operations:

(1) "Replace R_2 by $-2R_1 + R_2$," (2) "Replace R_3 by $3R_1 + R_3$," (3) "Replace R_3 by $-4R_2 + R_3$"

These operations yield the following, where the triangular form is U:

$$A \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & -4 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = U \quad \text{and} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 4 & 1 \end{bmatrix}$$

The entries $2, -3, 4$ in L are the negatives of the multipliers $-2, 3, -4$ in the above row operations. (As a check, multiply L and U to verify $A = LU$.)

3.42. Let A be the matrix in Problem 3.41. Find X_1, X_2, X_3, where X_i is the solution of $AX = B_i$ for (a) $B_1 = (1, 1, 1)$, (b) $B_2 = B_1 + X_1$, (c) $B_3 = B_2 + X_2$.

(a) Find $L^{-1}B_1$ by applying the row operations (1), (2), and then (3) in Problem 3.41 to B_1:

$$B_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \xrightarrow{\text{(1) and (2)}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \xrightarrow{\text{(3)}} \begin{bmatrix} 1 \\ -1 \\ 8 \end{bmatrix}$$

Solve $UX = B$ for $B = (1, -1, 8)$ by back-substitution to obtain $X_1 = (-25, 9, 8)$.

(b) First find $B_2 = B_1 + X_1 = (1, 1, 1) + (-25, 9, 8) = (-24, 10, 9)$. Then as above

$$B_2 = [-24, 10, 9]^T \xrightarrow{\text{(1) and (2)}} [-24, 58, -63]^T \xrightarrow{\text{(3)}} [-24, 58, -295]^T$$

Solve $UX = B$ for $B = (-24, 58, -295)$ by back-substitution to obtain $X_2 = (943, -353, -295)$.

(c) First find $B_3 = B_2 + X_2 = (-24, 10, 9) + (943, -353, -295) = (919, -343, -286)$. Then, as above

$$B_3 = [943, -353, -295]^T \xrightarrow{\text{(1) and (2)}} [919, -2181, 2671]^T \xrightarrow{\text{(3)}} [919, -2181, 11\,395]^T$$

Solve $UX = B$ for $B = (919, -2181, 11\,395)$ by back-substitution to obtain

$$X_3 = (-37\,628, 13\,576, 11\,395).$$

Miscellaneous Problems

3.43. Let L be a linear combination of the m equations in n unknowns in the system (3.2). Say L is the equation

$$(c_1a_{11} + \cdots + c_ma_{m1})x_1 + \cdots + (c_1a_{1n} + \cdots + c_ma_{mn})x_n = c_1b_1 + \cdots + c_mb_m \tag{1}$$

Show that any solution of the system (3.2) is also a solution of L.

Let $u = (k_1, \ldots, k_n)$ be a solution of (3.2). Then

$$a_{i1}k_1 + a_{i2}k_2 + \cdots + a_{in}k_n = b_i \qquad (i = 1, 2, \ldots, m) \tag{2}$$

Substituting u in the left-hand side of (1) and using (2), we get

$$\begin{aligned}(c_1a_{11} &+ \cdots + c_ma_{m1})k_1 \;+\; \cdots \;+\; (c_1a_{1n} + \cdots + c_ma_{mn})k_n \\ &= c_1(a_{11}k_1 + \cdots + a_{1n}k_n) \;+\; \cdots \;+\; c_m(a_{m1}k_1 + \cdots + a_{mn}k_n) \\ &= c_1b_1 + \cdots + c_mb_m\end{aligned}$$

This is the right-hand side of (1); hence, u is a solution of (1).

3.44. Suppose a system $\mathscr{M}$ of linear equations is obtained from a system $\mathscr{L}$ by applying an elementary operation (page 64). Show that $\mathscr{M}$ and $\mathscr{L}$ have the same solutions.

Each equation L in $\mathscr{M}$ is a linear combination of equations in $\mathscr{L}$. Hence, by Problem 3.43, any solution of $\mathscr{L}$ will also be a solution of $\mathscr{M}$. On the other hand, each elementary operation has an inverse elementary operation, so $\mathscr{L}$ can be obtained from $\mathscr{M}$ by an elementary operation. This means that any solution of $\mathscr{M}$ is a solution of $\mathscr{L}$. Thus, $\mathscr{L}$ and $\mathscr{M}$ have the same solutions.

3.45. Prove Theorem 3.4: Suppose a system $\mathscr{M}$ of linear equations is obtained from a system $\mathscr{L}$ by a sequence of elementary operations. Then $\mathscr{M}$ and $\mathscr{L}$ have the same solutions.

Each step of the sequence does not change the solution set (Problem 3.44). Thus, the original system $\mathscr{L}$ and the final system $\mathscr{M}$ (and any system in between) have the same solutions.

3.46. A system $\mathscr{L}$ of linear equations is said to be *consistent* if no linear combination of its equations is a degenerate equation L with a nonzero constant. Show that $\mathscr{L}$ is consistent if and only if $\mathscr{L}$ is reducible to echelon form.

Suppose $\mathscr{L}$ is reducible to echelon form. Then $\mathscr{L}$ has a solution, which must also be a solution of every linear combination of its equations. Thus, L, which has no solution, cannot be a linear combination of the equations in $\mathscr{L}$. Thus, $\mathscr{L}$ is consistent.

On the other hand, suppose $\mathscr{L}$ is not reducible to echelon form. Then, in the reduction process, it must yield a degenerate equation L with a nonzero constant, which is a linear combination of the equations in $\mathscr{L}$. Therefore, $\mathscr{L}$ is not consistent; that is, $\mathscr{L}$ is inconsistent.

3.47. Suppose u and v are distinct vectors. Show that, for distinct scalars k, the vectors $u + k(u - v)$ are distinct.

Suppose $u + k_1(u - v) = u + k_2(u - v)$. We need only show that $k_1 = k_2$. We have

$$k_1(u - v) = k_2(u - v), \qquad \text{and so} \qquad (k_1 - k_2)(u - v) = 0$$

Because u and v are distinct, $u - v \neq 0$. Hence, $k_1 - k_2 = 0$, and so $k_1 = k_2$.

3.48. Suppose AB is defined. Prove

(a) Suppose A has a zero row. Then AB has a zero row.

(b) Suppose B has a zero column. Then AB has a zero column.

(a) Let R_i be the zero row of A, and $C_1, \ldots, C_n$ the columns of B. Then the ith row of AB is

$$(R_iC_1, R_iC_2, \ldots, R_iC_n) = (0, 0, 0, \ldots, 0)$$

(b) B^T has a zero row, and so $B^TA^T = (AB)^T$ has a zero row. Hence, AB has a zero column.

SUPPLEMENTARY PROBLEMS

Linear Equations, 2 × 2 Systems

3.49. Determine whether each of the following systems is linear:

(a) $3x - 4y + 2yz = 8$, (b) $ex + 3y = \pi$, (c) $2x - 3y + kz = 4$

3.50. Solve (a) $\pi x = 2$, (b) $3x + 2 = 5x + 7 - 2x$, (c) $6x + 2 - 4x = 5 + 2x - 3$

3.51. Solve each of the following systems:

(a) $\begin{aligned} 2x + 3y &= 1 \\ 5x + 7y &= 3 \end{aligned}$ (b) $\begin{aligned} 4x - 2y &= 5 \\ -6x + 3y &= 1 \end{aligned}$ (c) $\begin{aligned} 2x - 4 &= 3y \\ 5y - x &= 5 \end{aligned}$ (d) $\begin{aligned} 2x - 4y &= 10 \\ 3x - 6y &= 15 \end{aligned}$

3.52. Consider each of the following systems in unknowns x and y:

(a) $\begin{aligned} x - ay &= 1 \\ ax - 4y &= b \end{aligned}$ (b) $\begin{aligned} ax + 3y &= 2 \\ 12x + ay &= b \end{aligned}$ (c) $\begin{aligned} x + ay &= 3 \\ 2x + 5y &= b \end{aligned}$

For which values of a does each system have a unique solution, and for which pairs of values (a, b) does each system have more than one solution?

General Systems of Linear Equations

3.53. Solve

(a) $\begin{aligned} x + y + 2z &= 4 \\ 2x + 3y + 6z &= 10 \\ 3x + 6y + 10z &= 17 \end{aligned}$ (b) $\begin{aligned} x - 2y + 3z &= 2 \\ 2x - 3y + 8z &= 7 \\ 3x - 4y + 13z &= 8 \end{aligned}$ (c) $\begin{aligned} x + 2y + 3z &= 3 \\ 2x + 3y + 8z &= 4 \\ 5x + 8y + 19z &= 11 \end{aligned}$

3.54. Solve

(a) $\begin{aligned} x - 2y &= 5 \\ 2x + 3y &= 3 \\ 3x + 2y &= 7 \end{aligned}$ (b) $\begin{aligned} x + 2y - 3z + 2t &= 2 \\ 2x + 5y - 8z + 6t &= 5 \\ 3x + 4y - 5z + 2t &= 4 \end{aligned}$ (c) $\begin{aligned} x + 2y + 4z - 5t &= 3 \\ 3x - y + 5z + 2t &= 4 \\ 5x - 4y + 6z + 9t &= 2 \end{aligned}$

3.55. Solve

(a) $\begin{aligned} 2x - y - 4z &= 2 \\ 4x - 2y - 6z &= 5 \\ 6x - 3y - 8z &= 8 \end{aligned}$ (b) $\begin{aligned} x + 2y - z + 3t &= 3 \\ 2x + 4y + 4z + 3t &= 9 \\ 3x + 6y - z + 8t &= 10 \end{aligned}$

3.56. Consider each of the following systems in unknowns x, y, z:

(a) $\begin{aligned} x - 2y &= 1 \\ x - y + az &= 2 \\ ay + 9z &= b \end{aligned}$ (b) $\begin{aligned} x + 2y + 2z &= 1 \\ x + ay + 3z &= 3 \\ x + 11y + az &= b \end{aligned}$ (c) $\begin{aligned} x + y + az &= 1 \\ x + ay + z &= 4 \\ ax + y + z &= b \end{aligned}$

For which values of a does the system have a unique solution, and for which pairs of values (a, b) does the system have more than one solution? The value of b does not have any effect on whether the system has a unique solution. Why?

Linear Combinations, Homogeneous Systems

3.57. Write v as a linear combination of u_1, u_2, u_3, where

(a) $v = (4, -9, 2)$, $u_1 = (1, 2, -1)$, $u_2 = (1, 4, 2)$, $u_3 = (1, -3, 2)$;
(b) $v = (1, 3, 2)$, $u_1 = (1, 2, 1)$, $u_2 = (2, 6, 5)$, $u_3 = (1, 7, 8)$;
(c) $v = (1, 4, 6)$, $u_1 = (1, 1, 2)$, $u_2 = (2, 3, 5)$, $u_3 = (3, 5, 8)$.

3.58. Let $u_1 = (1, 1, 2)$, $u_2 = (1, 3, -2)$, $u_3 = (4, -2, -1)$ in $\mathbf{R}^3$. Show that u_1, u_2, u_3 are orthogonal, and write v as a linear combination of u_1, u_2, u_3, where (a) $v = (5, -5, 9)$, (b) $v = (1, -3, 3)$, (c) $v = (1, 1, 1)$. (*Hint:* Use Fourier coefficients.)

3.59. Find the dimension and a basis of the general solution W of each of the following homogeneous systems:

(a) $\begin{aligned} x - y + 2z &= 0 \\ 2x + y + z &= 0 \\ 5x + y + 4z &= 0 \end{aligned}$ (b) $\begin{aligned} x + 2y - 3z &= 0 \\ 2x + 5y + 2z &= 0 \\ 3x - y - 4z &= 0 \end{aligned}$ (c) $\begin{aligned} x + 2y + 3z + t &= 0 \\ 2x + 4y + 7z + 4t &= 0 \\ 3x + 6y + 10z + 5t &= 0 \end{aligned}$

3.60. Find the dimension and a basis of the general solution W of each of the following systems:

(a) $\begin{aligned} x_1 + 3x_2 + 2x_3 - x_4 - x_5 &= 0 \\ 2x_1 + 6x_2 + 5x_3 + x_4 - x_5 &= 0 \\ 5x_1 + 15x_2 + 12x_3 + x_4 - 3x_5 &= 0 \end{aligned}$ (b) $\begin{aligned} 2x_1 - 4x_2 + 3x_3 - x_4 + 2x_5 &= 0 \\ 3x_1 - 6x_2 + 5x_3 - 2x_4 + 4x_5 &= 0 \\ 5x_1 - 10x_2 + 7x_3 - 3x_4 + 18x_5 &= 0 \end{aligned}$

Echelon Matrices, Row Canonical Form

3.61. Reduce each of the following matrices to echelon form and then to row canonical form:

(a) $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 4 & 9 \\ 1 & 5 & 12 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 2 & -1 & 2 & 1 \\ 2 & 4 & 1 & -2 & 5 \\ 3 & 6 & 3 & -7 & 7 \end{bmatrix}$, (c) $\begin{bmatrix} 2 & 4 & 2 & -2 & 5 & 1 \\ 3 & 6 & 2 & 2 & 0 & 4 \\ 4 & 8 & 2 & 6 & -5 & 7 \end{bmatrix}$

3.62. Reduce each of the following matrices to echelon form and then to row canonical form:

(a) $\begin{bmatrix} 1 & 2 & 1 & 2 & 1 & 2 \\ 2 & 4 & 3 & 5 & 5 & 7 \\ 3 & 6 & 4 & 9 & 10 & 11 \\ 1 & 2 & 4 & 3 & 6 & 9 \end{bmatrix}$, (b) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 3 & 8 & 12 \\ 0 & 0 & 4 & 6 \\ 0 & 2 & 7 & 10 \end{bmatrix}$, (c) $\begin{bmatrix} 1 & 3 & 1 & 3 \\ 2 & 8 & 5 & 10 \\ 1 & 7 & 7 & 11 \\ 3 & 11 & 7 & 15 \end{bmatrix}$

3.63. Using only 0's and 1's, list all possible 2×2 matrices in row canonical form.

3.64. Using only 0's and 1's, find the number n of possible 3×3 matrices in row canonical form.

Elementary Matrices, Applications

3.65. Let e_1, e_2, e_3 denote, respectively, the following elementary row operations:

"Interchange R_2 and R_3," "Replace R_2 by $3R_2$," "Replace R_1 by $2R_3 + R_1$"

(a) Find the corresponding elementary matrices E_1, E_2, E_3.
(b) Find the inverse operations e_1^{-1}, e_2^{-1}, e_3^{-1}; their corresponding elementary matrices E'_1, E'_2, E'_3; and the relationship between them and E_1, E_2, E_3.
(c) Describe the corresponding elementary column operations f_1, f_2, f_3.
(d) Find elementary matrices F_1, F_2, F_3 corresponding to f_1, f_2, f_3, and the relationship between them and E_1, E_2, E_3.

3.66. Express each of the following matrices as a product of elementary matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}, \qquad C = \begin{bmatrix} 2 & 6 \\ -3 & -7 \end{bmatrix}, \qquad D = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 3 & 8 & 7 \end{bmatrix}$$

3.67. Find the inverse of each of the following matrices (if it exists):

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 2 & -3 & 1 \\ 3 & -4 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 6 & 1 \\ 3 & 10 & -1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 8 & -3 \\ 1 & 7 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 2 & 1 & -1 \\ 5 & 2 & -3 \\ 0 & 2 & 1 \end{bmatrix}$$

3.68. Find the inverse of each of the following $n \times n$ matrices:

(a) A has 1's on the diagonal and *superdiagonal* (entries directly above the diagonal) and 0's elsewhere.

(b) B has 1's on and above the diagonal, and 0's below the diagonal.

Lu Factorization

3.69. Find the LU factorization of each of the following matrices:

(a) $\begin{bmatrix} 1 & -1 & -1 \\ 3 & -4 & -2 \\ 2 & -3 & -2 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 3 & -1 \\ 2 & 5 & 1 \\ 3 & 4 & 2 \end{bmatrix}$, (c) $\begin{bmatrix} 2 & 3 & 6 \\ 4 & 7 & 9 \\ 3 & 5 & 4 \end{bmatrix}$, (d) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 10 \end{bmatrix}$

3.70. Let A be the matrix in Problem 3.69(a). Find X_1, X_2, X_3, X_4, where

(a) X_1 is the solution of $AX = B_1$, where $B_1 = (1, 1, 1)^T$.

(b) For $k > 1$, X_k is the solution of $AX = B_k$, where $B_k = B_{k-1} + X_{k-1}$.

3.71. Let B be the matrix in Problem 3.69(b). Find the LDU factorization of B.

Miscellaneous Problems

3.72. Consider the following systems in unknowns x and y:

$$\text{(a)} \quad \begin{aligned} ax + by &= 1 \\ cx + dy &= 0 \end{aligned} \qquad\qquad \text{(b)} \quad \begin{aligned} ax + by &= 0 \\ cx + dy &= 1 \end{aligned}$$

Suppose $D = ad - bc \neq 0$. Show that each system has the unique solution:

(a) $x = d/D, y = -c/D$, (b) $x = -b/D, y = a/D$.

3.73. Find the inverse of the row operation "Replace R_i by $kR_j + k'R_i$ $(k' \neq 0)$."

3.74. Prove that deleting the last column of an echelon form (respectively, the row canonical form) of an augmented matrix $M = [A, B]$ yields an echelon form (respectively, the row canonical form) of A.

3.75. Let e be an elementary row operation and E its elementary matrix, and let f be the corresponding elementary column operation and F its elementary matrix. Prove

(a) $f(A) = (e(A^T))^T$, (b) $F = E^T$, (c) $f(A) = AF$.

3.76. Matrix A is *equivalent* to matrix B, written $A \approx B$, if there exist nonsingular matrices P and Q such that $B = PAQ$. Prove that $\approx$ is an *equivalence* relation; that is,

(a) $A \approx A$, (b) If $A \approx B$, then $B \approx A$, (c) If $A \approx B$ and $B \approx C$, then $A \approx C$.

ANSWERS TO SUPPLEMENTARY PROBLEMS

Notation: $A = [R_1;\quad R_2;\quad \ldots]$ denotes the matrix A with rows $R_1, R_2, \ldots$. The elements in each row are separated by commas (which may be omitted with single digits), the rows are separated by semicolons, and 0 denotes a zero row. For example,

$$A = [1,2,3,4;\quad 5,-6,7,-8;\quad 0] = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & -6 & 7 & -8 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

3.49. (a) no, (b) yes, (c) linear in x, y, z, not linear in x, y, z, k

3.50. (a) $x = 2/\pi$, (b) no solution, (c) every scalar k is a solution

3.51. (a) $(2, -1)$, (b) no solution, (c) $(5, 2)$, (d) $(5 - 2a,\ a)$

3.52. (a) $a \neq \pm 2$, $(2, 2)$, $(-2, -2)$, (b) $a \neq \pm 6$, $(6, 4)$, $(-6, -4)$, (c) $a \neq \frac{5}{2}$, $(\frac{5}{2}, 6)$

3.53. (a) $(2, 1, \frac{1}{2})$, (b) no solution, (c) $u = (-7a - 1,\ 2a + 2,\ a)$.

3.54. (a) $(3, -1)$, (b) $u = (-a + 2b,\ 1 + 2a - 2b,\ a,\ b)$, (c) no solution

3.55. (a) $u = (\frac{1}{2}a + 2,\quad a,\quad \frac{1}{2})$, (b) $u = (\frac{1}{2}(7 - 5b - 4a),\quad a,\ \frac{1}{2}(1 + b),\ b)$

3.56. (a) $a \neq \pm 3$, $(3, 3)$, $(-3, -3)$, (b) $a \neq 5$ and $a \neq -1$, $(5, 7)$, $(-1, -5)$,
(c) $a \neq 1$ and $a \neq -2$, $(-2, 5)$

3.57. (a) $2, -1, 3$, (b) $6, -3, 1$, (c) not possible

3.58. (a) $3, -2, 1$, (b) $\frac{2}{3}, -1, \frac{1}{3}$, (c) $\frac{2}{3}, \frac{1}{7}, \frac{1}{21}$

3.59. (a) $\dim W = 1$, $u_1 = (-1, 1, 1)$, (b) $\dim W = 0$, no basis,
(c) $\dim W = 2$, $u_1 = (-2, 1, 0, 0)$, $u_2 = (5, 0, -2, 1)$

3.60. (a) $\dim W = 3$, $u_1 = (-3, 1, 0, 0, 0)$, $u_2 = (7, 0, -3, 1, 0)$, $u_3 = (3, 0, -1, 0, 1)$,
(b) $\dim W = 2$, $u_1 = (2, 1, 0, 0, 0)$, $u_2 = (5, 0, -5, -3, 1)$

3.61. (a) $[1, 0, -\frac{1}{2};\quad 0, 1, \frac{5}{2};\quad 0]$, (b) $[1, 2, 0, 0, 2;\quad 0, 0, 1, 0, 5;\quad 0, 0, 0, 1, 2]$,
(c) $[1, 2, 0, 4, -5, 3;\quad 0, 0, 1, -5, \frac{15}{2}, -\frac{5}{2};\quad 0]$

3.62. (a) $[1, 2, 0, 0, -4, -2;\quad 0, 0, 1, 0, 1, 2;\quad 0, 0, 0, 1, 2, 1;\quad 0]$,
(b) $[0, 1, 0, 0;\quad 0, 0, 1, 0;\quad 0, 0, 0, 1;\quad 0]$, (c) $[1, 0, 0, 4;\quad 0, 1, 0, -1;\quad 0, 0, 1, 2;\quad 0]$

3.63. 5: $[1, 0;\quad 0, 1]$, $[1, 1;\quad 0, 0]$, $[1, 0;\quad 0, 0]$, $[0, 1;\quad 0, 0]$, 0

3.64. 16

3.65. (a) $[1, 0, 0;\quad 0, 0, 1;\quad 0, 1, 0]$, $[1, 0, 0;\quad 0, 3, 0;\quad 0, 0, 1]$, $[1, 0, 2;\quad 0, 1, 0;\quad 0, 0, 1]$,
(b) $R_2 \leftrightarrow R_3$; $\frac{1}{3}R_2 \rightarrow R_2$; $-2R_3 + R_1 \rightarrow R_1$; each $E_i' = E_i^{-1}$,
(c) $C_2 \leftrightarrow C_3, 3C_2 \rightarrow C_2, 2C_3 + C_1 \rightarrow C_1$, (d) each $F_i = E_i^T$.

3.66. $A = [1, 0;\quad 3, 1][1, 0;\quad 0, -2][1, 2;\quad 0, 1]$, B is not invertible,
$C = [1, 0;\quad -\frac{3}{2}, 1][1, 0;\quad 0, 2][1, 6;\quad 0, 1][2, 0;\quad 0, 1]$,
$D = [100;\quad 010;\quad 301][100;\quad 010;\quad 021][100;\quad 013;\quad 001][120;\quad 010;\quad 001]$

3.67. $A^{-1} = [-8, 12, -5;\quad -5, 7, -3;\quad 1, -2, 1]$, B has no inverse,
$C^{-1} = [\frac{29}{2}, -\frac{17}{2}, \frac{7}{2};\quad -\frac{5}{2}, \frac{3}{2}, -\frac{1}{2};\quad 3, -2, 1]$, $D^{-1} = [8, -3, -1;\quad -5, 2, 1;\quad 10, -4, -1]$

3.68. $A^{-1} = [1, -1, 1, -1, \ldots; \quad 0, 1, -1, 1, -1, \ldots; \quad 0, 0, 1, -1, 1, -1, 1, \ldots; \quad \ldots; \quad \ldots; \quad 0, \ldots 0, 1]$
B^{-1} has 1's on diagonal, -1's on superdiagonal, and 0's elsewhere.

3.69. (a) $[100; \quad 310; \quad 211][1, -1, -1; \quad 0, -1, 1; \quad 0, 0, -1]$,
(b) $[100; \quad 210; \quad 351][1, 3, -1; \quad 0, -1, 3; \quad 0, 0, -10]$,
(c) $[100; \quad 210; \quad \frac{3}{2}, \frac{1}{2}, 1][2, 3, 6; \quad 0, 1, -3; \quad 0, 0, -\frac{7}{2}]$,
(d) There is no LU decomposition.

3.70. $X_1 = [1, 1, -1]^T, \quad B_2 = [2, 2, 0]^T, \quad X_2 = [6, 4, 0]^T, \quad B_3 = [8, 6, 0]^T, \quad X_3 = [22, 16, -2]^T,$
$B_4 = [30, 22, -2]^T, \quad X_4 = [86, 62, -6]^T$

3.71. $B = [100; \quad 210; \quad 351] \operatorname{diag}(1, -1, -10) [1, 3, -1; \quad 0, 1, 3; \quad 0, 0, 1]$

3.73. Replace R_i by $-kR_j + (1/k')R_i$.

3.75. (c) $f(A) = (e(A^T))^T = (EA^T)^T = (A^T)^T E^T = AF$

3.76. (a) $A = IAI$. (b) If $A = PBQ$, then $B = P^{-1}AQ^{-1}$.
(c) If $A = PBQ$ and $B = P'CQ'$, then $A = (PP')C(Q'Q)$.